Niveau(x) d'enseignement : 7 Numéro EDUQ : 9470362
Centre d'intérêt : 540

Howe, Robert.

Croyances et pratiques en évaluation des apprentissages : étude des croyances et des pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages / Robert Howe, Louise Ménard. -- Laval : Collège Montmorency, 1993.
xx, 404 p. ; 28 cm.

Résumé : p. xix-xx. Code : 1532-0412. Bibliogr. : p. 265-276.

ISBN 2-89058-029-6

Vedette(s)-secondaire(s) : Ménard, Louise * Collège Montmorency

Descripteur(s) : Evaluation de l'apprentissage * Opinion de l'enseignant * Attitude de l'enseignant * Pratique pédagogique * Compétence de l'enseignant * Formation des enseignants * Perfectionnement des enseignants * Questionnaire * Collégial (Ordre d'enseignement)

Identificateur(s) de sujet : Croyances de l'enseignant

Cette recherche s'intéresse aux croyances des enseignants du collégial en matière d'évaluation des apprentissages de façon à contribuer à l'amélioration de leurs pratiques à l'égard de la mesure et de l'évaluation et de pouvoir dégager des pistes de perfectionnement possibles. A cette fin, des données quantitatives et qualitatives sont recueillies auprès de 1690 enseignants, répartis dans les 47 cégeps, et interrogés par questionnaire. Il ressort de l'analyse des données les constatations suivantes : la méconnaissance de l'évaluation formative; l'utilisation de l'évaluation à des fin détournées ainsi que la pollution des notes versées au bulletin; le caractère facultatif de la formation des enseignants en mesure et évaluation. Questionnaire à l'annexe 1.

Type de document :

COLLÈGE MONTMORENCY

CROYANCES ET PRATIQUES EN ÉVALUATION DES APPRENTISSAGES

ROBERT HOWE, conseiller pédagogique
LOUISE MÉNARD, conseillère pédagogique

JUIN 1993

147362

Croyances et pratiques en évaluation des apprentissages

Étude des croyances et des pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages

Robert Howe, Conseiller pédagogique
Louise Ménard, Conseillère pédagogique

Collège Montmorency
Juin 1993

Cette recherche a été subventionnée par le ministère de l'Enseignement supérieur et de la Science dans le cadre du programme d'aide à la recherche sur l'enseignement et l'apprentissage.

Le contenu du présent rapport n'engage que la responsabilité du collège et de ses auteurs.

On peut obtenir des exemplaires supplémentaires de ce rapport de recherche auprès de la Direction des services pédagogiques du collège Montmorency, 475 boulevard de l'Avenir, Laval, Québec, H7N 5H9

Maquette de page couverture: Michel Belhumeur, Service des communications, collège Montmorency

Impression: Collège Montmorency
Dépôt légal: second trimestre 1993
Bibliothèque nationale du Québec
ISBN: 2-89058-029-6
Collège Montmorency
Juin 1993

Ce rapport est dédié à
nos enfants qui ont su ... patienter
et à
l'abbé Roger Rinfret

REMERCIEMENTS

Nous tenons à remercier toutes les personnes qui, grâce à leur collaboration, ont rendu possible la réalisation de cette recherche. Nous voulons tout particulièrement remercier:

- Les répondants de chaque collège et leur secrétaire qui ont collaboré à la distribution de nos questionnaires auprès de tous les professeurs désignés dans notre échantillon;

- Les consultants experts qui ont validé notre questionnaire ainsi que le domaine de la recherche;

- Monsieur Philippe Ricard, du Bureau d'études sociographiques, qui a assuré la mise en page du questionnaire, la saisie et le traitement des données ainsi que le counseling méthodologique pour le traitement et l'analyse des données;

- Madame France Fontaine, adjointe à la Doyenne des études de premier cycle à l'U.Q.A.M., pour ses conseils méthodologiques et conceptuels en évaluation des apprentissages;

- Mesdames Lucie Garneau et Brigitte Bouchard pour le traitement de texte et la mise en page;

- Monsieur Jean-Pierre Regnault pour ses conseils de mise en page et d'édition;

- La direction des Services éducatifs des trois collèges qui nous ont permis d'y rencontrer des groupes réacteurs;

- ...et, bien entendu, tous les professeurs qui ont consenti à répondre au questionnaire ou qui nous ont rencontrés en entrevue individuelle ou en groupe.

Ce rapport de recherche a été produit entièrement sur Microsoft Word 4.0 que nous avons appris à utiliser, en autodidactes, grâce à l'excellent manuel de Jean-Marie Tremblay, professeur de sociologie au cégep de Chicoutimi: Microsoft Word 4.0 pour le Macintosh d'Apple, version française, 1990, 340 pages.

Les logiciels Statview 1.04 (statistiques) et Pro-Cite 2.0 (bibliographie) ont été mis à contribution.

TABLE DES MATIÈRES

REMERCIEMENTS V

TABLE DES MATIÈRES VII

LISTE DES TABLEAUX XIII

RÉSUMÉ XIX

INTRODUCTION 1

1. LE PROBLÈME 7
1.1 Définition du problème et objet de la recherche 9
1.2 Les objectifs de la recherche 13
1.3 Le type de recherche privilégié 13
1.4 Les limites de la recherche 14
1.4.1 Limites du domaine 14
1.4.2 Limites méthodologiques 15
1.5 Définition des termes 17

2. ÉTAT DE LA QUESTION 19
2.1 La question des croyances 21
2.1.1 Le concept de croyances 21
2.1.2 La relation entre les attitudes, les croyances et les pratiques 22
2.1.3 Genèse des croyances 23
2.1.4 Altérabilité des croyances 25
2.1.5 Comprendre les attitudes pour guider le changement: Le modèle de Pryor 28
2.1.6 Genèse des attitudes et cycle d'évaluation 29
2.1.7 Persistance des croyances 31
2.1.8 Profession enseignante et altérabilité des croyances 32
2.1.9 Les croyances et les pratiques des enseignants en évaluation des apprentissages 33

2.2 La question des pratiques 33
2.2.1 Apprendre à évaluer 33
- Formation académique 34
- Apprentissage expérientiel et folklore 35
2.2.2 Les pratiques en évaluation des apprentissages 36
- Autonomie dans l'évaluation 36
- Instrumentation 38
- Notation 40
2.2.3 Définir les compétences en évaluation 44
- Les standards de compétence 45
2.2.4 Altérabilité des pratiques 46

3. ÉVALUATION, ENSEIGNEMENT ET APPRENTISSAGE 51
3.1 Le paradigme de l'évaluation des apprentissages 53
3.1.1 Évolution du paradigme de l'évaluation 53
- L'ancien paradigme en évaluation ou le paradigme du "Test Theory" 53
- Le nouveau paradigme en évaluation ou l'évaluation intégrée au processus d'enseignement et d'apprentissage 56
3.1.2 L'évaluation, une aide à l'apprentissage 58
3.1.3 De l'évaluation à l'assessment: Une question d'attitude 59
3.2 Quelques conséquences de l'évaluation 61
3.2.1 L'évaluation et l'enseignement 61
- Les impacts sur la notation 62
- Les impacts sur la tâche 63
- Les impacts sur la pédagogie 65
3.2.2 L'évaluation et l'apprentissage 65
3.3 De la docimologie à la mesure et évaluation 67

4. LA MÉTHODOLOGIE 69
4.1 Aperçu général 71
4.2 Élaboration du questionnaire 72
4.2.1 Extraction des énoncés 72
4.2.2 Le choix d'une typologie 73
4.2.3 Le système de catégories de Stiggins 78
4.2.4 Élagage et ajout d'énoncés 81
4.2.5 Des entrevues de pré-consultation 82
4.3 Validation et mise en forme du questionnaire 84
4.3.1 Pré-validation 84
4.3.2 Mise en page 84
4.3.3 Validation de la pénultième version 85
4.4 L'échantillon 85
4.5 La distribution des questionnaires 90
4.6 Les groupes réacteurs 91

5. LES RÉSULTATS DE LA RECHERCHE 93
5.1 Composition du chapitre 95
5.1.1 Contenu du chapitre 95
5.1.2 Composition des tableaux 95
- Les tableaux de fréquences de l'échantillon total 95
- Les tableaux d'analyses stratifiées 96
5.1.3 Structure du chapitre 98
5.2 Présentation des connées 100
5.2.1 Première catégorie de Stiggins: L'utilisation de l'évaluation en classe 100
- Les croyances relatives à l'utilisation de l'évaluation en classe 100
- Les pratiques relatives à l'utilisation de l'évaluation en classe 104
5.2.2 Deuxième catégorie de Stiggins: Les cibles de l'évaluation 108
- Les croyances relatives aux cibles de l'évaluation 108
- Les pratiques relatives aux cibles de l'évaluation 110

5.2.3 Troisième catégorie de Stiggins: Les qualités de l'évaluation 114
- Les croyances relatives aux qualités de l'évaluation 114
- Les pratiques relatives aux qualités de l'évaluation 118
5.2.4 Quatrième catégorie de Stiggins: Les instruments d'évaluation 128
- Les croyances relatives aux instruments 128
- Les pratiques relatives aux instruments d'évaluation 133
5.2.5 Cinquième catégorie de Stiggins: Les dimensions interpersonnelles de l'évaluation 140
- Les croyances relatives aux dimensions interpersonnelles de l'évaluation 140
- Les pratiques relatives aux dimensions interpersonnelles de l'évaluation 142
5.2.6 Sixième catégorie de Stiggins: La rétroaction et la notation 146
- Les croyances relatives à la rétroaction et à la notation 146
- Les pratiques relatives à la rétroaction et à la notation 149
5.2.7 Les questions portant sur la formation et le perfectionnement 152

6. ANALYSE DES RÉSULTATS 161
6.1 Considérations d'analyse 163
6.1.1 Les groupes réacteurs: Attitudes générales 164
6.1.2 Les sources de variation dans les profils de pratiques 165
6.1.3 Jugements de valeur 167
6.1.4 Structure du chapitre 167
6.1.5 Analyses statistiques 168
6.1.6 À propos des recommandations 170
6.2 Analyse des résultats 171
6.2.1 Le premier champ de compétence: L'utilisation de l'évaluation en classe 171
- Analyse 171
- Analyse factorielle 181
6.2.2 Le second champ de compétence: Les cibles de l'évaluation 183
- Analyse 183
- Analyse factorielle 185

6.2.3 Le troisième champ de compétence: Les qualités de l'évaluation 188
- Analyse 188
- Analyse factorielle 198
6.2.4 Le quatrième champ de compétence: Les instruments d'évaluation 202
- Analyse 202
6.2.5 Le cinquième champ de compétence: Les dimensions interpersonnelles de l'évaluation 207
- Analyse 207
- Analyse factorielle 211
6.2.6 Le sixième champ de compétence: La rétroaction et la notation 218
- Analyse 218
- Analyse factorielle 222
6.3 La formation en mesure et évaluation: Analyse et recommandations 225
6.3.1 La prépation des enseignants en mesure et évaluation des apprentissages 225
6.3.2 L'autoévaluation des compétences en évaluation des apprentissages 226
6.3.3 La formation des professeurs 228
- Les cibles de la formation 229
- Les lieux de la formation 236
- La nature de la formation 238
6.3.4 Cadres et conseillers pédagogiques: Perfectionnement en évaluation des apprentissages 244
6.3.5 Les collaborateurs à ces formations 246
6.3.6 Services à offrir 247

CONCLUSION 249

BIBLIOGRAPHIE 263

ANNEXE 1: LE QUESTIONNAIRE 277

ANNEXE 2: VALIDATION DU QUESTIONNAIRE 295

ANNEXE 3: LE PROTOCOLE D'ENTREVUE 333

ANNEXE 4: POPULATION ET ÉCHANTILLON 339

ANNEXE 5: TABLEAUX D'ANALYSES STRATIFIÉES 345

ANNEXE 6: ANALYSES FACTORIELLES 379

ANNEXE 7: CATÉGORIES D'ÉNONCÉS DE CROYANCES ET DE PRATIQUES 395

LISTE DES TABLEAUX

CHAPITRE 4: LA MÉTHODOLOGIE 69

Tableau 1: Comparaison de quatre catégorisations du domaine de l'évaluation des apprentissages. 76

Tableau 2: Distribution des proportions attribuées à chaque strate 87

CHAPITRE 5: LES RÉSULTATS DE LA RECHERCHE 93

Tableau 3: Appartenance des énoncés du questionnaire à chacun des six champs de compétence 99

Tableau C1-a: Distribution des réponses aux énoncés de croyances portant sur l'utilisation de l'évaluation en classe 102

Tableau P1: Nombre d'heures, sur la base d'un cours de 45 heures, consacré à des activités d'évaluation 106

Tableau C2-a: Pourcentages des réponses aux énoncés de croyances portant sur les cibles de l'évaluation 110

Tableau P2-a: Proportions obtenues à la question 3: "Dans ce cours, en plus d'attribuer des points pour les objectifs liés à la matière enseignée, j'en accorde pour:" 113

Tableau P2-a: (suite) Qui fait les choix en matière de... 113

Tableau C3-a: Réponses aux énoncés de croyances portant sur les qualités de l'évaluation 117

Tableau P3-a: Les pratiques relatives aux qualités de l'évaluation. Réponses de l'échantillon total 121

Tableau P3-a: (suite) Les pratiques relatives aux qualités de l'évaluation. Réponses de l'échantillon total 124

Tableau C4-a: Pourcentages des réponses aux énoncés de croyances portant sur les instruments d'évaluation 130

Tableau P4-a:	Les approches utilisées pour évaluer les apprentissages, en ordre décroissant d'utilisation. Le pourcentage est constitué du cumul des "toujours" et "souvent" à la Q34: "Utilisez-vous les approches suivantes pour effectuer l'évaluation de l'apprentissage de vos étudiants?"	136
Tableau P4-a: (suite)	Palmarès des cinq (5) approches d'évaluation les plus invoquées, pour chacune des questions 35 à 40	138
Tableau P4-a: (suite)	Palmarès des cinq (5) approches d'évaluation formative les plus invoquées à la question 41	139
Tableau C5-a:	Pourcentages des réponses aux énoncés de croyances portant sur les dimensions interpersonnelles de l'évaluation	141
Tableau P5-a:	Pratiques d'intervention sur les notes. "Il m'arrive de..."	144
Tableau P5-a: (suite)	Pratiques d'intervention sur les notes. "Il m'arrive de..."	144
Tableau C6-a:	Pourcentages des réponses aux énoncés de croyances portant sur la rétroaction et la notation	147
Tableau P6-a:	Description des pratiques de rétroaction et de notation	150
Tableau Q65:	Niveaux de scolarité des répondants, par ordre décroissant.	159

CHAPITRE 6:	ANALYSE DES RÉSULTATS	161
Tableau 4:	Corrélations entre les réponses aux variables équivalentes des questions 3 et 45, 46.	209
ANNEXE 5:	TABLEAUX D'ANALYSES STRATIFIÉES	345
Tableau C1-b:	Croyances relatives à l'utilisation de l'évaluation en classe: Analyse des questions 6d) et 6j)	349
Tableau C2-b:	Croyances relatives aux cibles de l'évaluation	350

Tableau P2-b:	Analyse, par strate, des habiletés mesurées, selon la question 2: "Mes instruments d'évaluation visent la mesure de:"	351
Tableau P2-b: (suite)	Q2: Mes instruments d'évaluation visent la mesure de:	352
Tableau P2-b: (suite)	Analyse stratifiée des réponses aux questions 1 et 3	353
Tableau C3-b:	La validation des items	354
Tableau C3-b: (suite)	Critères de correction et représentativité	355
Tableau P3-b:	Les pratiques relatives aux qualités de l'évaluation	356
Tableau P3-b: (suite)	Pratiques relatives à la fréquence de l'évaluation	357
Tableau P3-b: (suite)	Pratiques de présentation de questions d'examen. Strates affichant des proportions de 3% et plus sur les réponses "Toujours", "Souvent" ou "Parfois"	358
Tableau P3-b: (suite)	Administration de l'évaluation: "Dans mes examens à développement..."	359
Tableau P3-b: (suite)	Pratiques de préparation des étudiants	360
Tableau P3-b: (suite)	Pratiques de correction et de notation	361
Tableau P3-b: (suite)	Pratiques d'amélioration a posteriori	362
Tableau C4-b:	Les croyances relatives aux instruments d'évaluation	363

Tableau C4-b: (suite)	Les croyances relatives aux instruments d'évaluation	364
Tableau P4-b:	Proportion de répondants qui disent utiliser l'examen ou le travail écrit pour effectuer l'évaluation de leurs étudiants	365
Tableau P4-b: (suite)	Dans mes examens, j'utilise les genres d'items suivants:	366
Tableau P4-b: (suite)	Distribution, dans l'ordre décroissant, de l'utilisation des diverses approches d'évaluation	367
Tableau P4-b: (suite)	Distribution, dans l'ordre décroissant, de l'utilisation des diverses approches d'évaluation	368
Tableau P4-b: (suite)	Distribution, dans l'ordre décroissant, de l'utilisation des diverses approches d'évaluation	369
Tableau P4-b: (suite)	Distribution, dans l'ordre décroissant, de l'utilisation des diverses approches d'évaluation	370
Tableau C5-b:	Les croyances relatives aux dimensions interpersonnelles de l'évaluation	371
Tableau P5-b:	Analyse, strate par strate, des interventions sur des notes finales des étudiants. Pourcentages intra-strates de réponses affirmatives aux questions 45 et 46. Présentation dans l'ordre décroissant des pratiques. "Il m'arrive de..."	372
Tableau P5-b: (suite)	Pratiques d'intervention sur les notes versées au bulletin, favorisant soit le groupe-cours, soit des individus	373
Tableau C6-b:	Analyse des croyances pour lesquelles il y a différence significative entre les strates	374
Tableau P6-b:	Les pratiques de rétroaction et de notation: Analyse stratifiée	375

Tableau Q55: Analyse stratifiée des sources de formation en évaluation des apprentissages 376

Tableau Q56: Autoévaluation des compétences en évaluation des apprentissages: "Comment jugez-vous vos compétences dans les domaines suivants?" 377

Tableau Q57: Proportions de répondants ayant indiqué qu'ils utilisent l'une ou plusieurs des sept formules pédagogiques ci-après. Ces formules pédagogiques sont présentées dans l'ordre décroissant des proportions observées dans l'échantillon 378

RÉSUMÉ

Dans les cégeps, la mesure et l'évaluation des apprentissages des étudiants relèvent entièrement et exclusivement des enseignants. Il est difficile d'enseigner; il est aussi très difficile d'évaluer. De nombreuses recherches et rapports décrivent et confirment les défis auxquels font face les professeurs à l'égard de l'évaluation des apprentissages.

Mais les enseignants des cégeps disposent de formations inégales, sinon aléatoires, en mesure et évaluation. Or, plusieurs auteurs décrivent l'immense influence des pratiques d'évaluation sur toutes les dimensions du processus d'enseignement et d'apprentissage. D'autres recherches, portant sur les attitudes, les croyances et les valeurs, démontrent qu'il faut s'intéresser d'abord aux croyances des enseignants si l'on veut amener quelque changement dans les pratiques.

Devant les enjeux pédagogiques de l'évaluation des apprentissages, la présente étude est fondée sur le désir de contribuer à l'amélioration des pratiques des enseignants à l'égard de la mesure et de l'évaluation. De nature descriptive, cette recherche se donne les trois objectifs suivants:

1. Identifier les croyances des enseignants des cégeps à l'égard de l'évaluation des apprentissages en salle de classe;

2. Décrire les pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages et ce, en situation régulière de salle de classe;

3. Dégager des pistes de perfectionnement qui prennent en compte les croyances et les pratiques des enseignants du collégial en matière d'évaluation des apprentissages en salle de classe.

Afin d'atteindre ces trois objectifs, des données quantitatives et qualitatives sont recueillies. Un échantillon aléatoire stratifié de 1,690 enseignants, répartis dans les 47 cégeps québécois, est interrogé par questionnaire. Les données quantitatives sont commentées et confirmées lors de rencontres de professeurs constitués en groupes réacteurs.

Des tableaux descriptifs de l'ensemble des répondants, des tableaux comparatifs par strates, des analyses factorielles et des commentaires extraits des groupes réacteurs permettent d'identifier des croyances et des pratiques qui illustrent certaines lacunes dont, tout particulièrement:

- la méconnaissance de l'évaluation formative;
- l'utilisation de l'évaluation à des fins détournées ainsi que la pollution des notes versées au bulletin;
- le caractère facultatif de la formation des professeurs en mesure et évaluation.

Le rapport de recherche propose plusieurs recommandations portant notamment sur le développement des compétences des professeurs ainsi que sur des services à offrir.

INTRODUCTION

Origine de cette recherche

Lorsqu'on dit qu'une recherche consiste à générer des savoirs nouveaux qui contribuent à répondre à des questions, on se surprend à constater que ces questions émergent souvent d'anecdotes, de faits divers qui, isolément, pourraient passer pour triviaux, mais, au cumul des ans, inspirent une énergie considérable lorsqu'un fait, en soi innocent, vient servir de déclencheur.

Un matin, c'était un 17 décembre, un professeur très respecté nous avoue, au vu de son air hagard, qu'il a corrigé les travaux de ses étudiants toute la nuit. Il a mis des heures à écrire des commentaires sur chacune des copies et pourra remettre ses notes demain, si tout va bien. Mais pourquoi, sachant que la session est terminée et que ses étudiants partent en vacances, a-t-il consacré tant d'effort à annoter les copies des étudiants de ses cinq groupes-cours? Parce que, répondit-il, il passera pour un professeur incompétent s'il ne le faisait pas.

Que les choix de cet enseignant soient fondés, dans les circonstances, sur une nécessité inhérente à la tâche ou sur des chimères, cette anecdote illustre une parcelle de l'impact de l'évaluation sur l'ensemble de son travail.

Pendant que cet enseignant affirme sur un air résigné qu'il doit "corriger" pendant des heures, d'autres s'interrogent sur des aspects éthiques de ce qu'ils font. Est-ce correct de donner des notes pour l'effort fourni? Tel collègue a-t-il le droit d'accorder 30 points pour les "devoirs" faits à la maison? Est-ce plus grave de laisser passer un étudiant qui aurait dû échouer ou de faire échouer quelqu'un qui mérite de passer?

Sur certains aspects, des questions cèdent la place à des affirmations catégoriques. Si on ne donne pas de notes pour l'évaluation formative, les étudiants n'étudieront pas. Les questions objectives sont à proscrire: elles ne font pas écrire et ne font pas réfléchir. Les questions à développement sont à proscrire: elles n'incitent pas à l'étude.

Et la lecture de certains plans de cours réserve parfois quelques surprises. "Les notes entre 56 et 59% seront arrondies à 60%" a-t-on lu dans un collège. "La note finale contiendra 10% pour l'effort fourni et 10% pour la qualité de la participation en classe".

De telles anecdotes pourraient faire l'objet d'un répertoire volumineux. Elles illustrent les très nombreuses questions que soulève, chez les professeurs, la fonction d'évaluation. Mais elles sont troublantes. Elles montrent que les professeurs évaluent les apprentissages sans nécessairement avoir la formation suffisante pour résoudre de nombreux problèmes techniques ou éthiques inhérents à cette tâche. Elles attirent l'attention sur la réalité de la formation des enseignants des collèges en mesure et évaluation. Elles suggèrent aussi que des pratiques d'évaluation sont mises en oeuvre dans nos collèges sans qu'elles soient toutes pertinentes.

Alors que l'évaluation a des impacts importants sur le fardeau de la tâche, qu'elle sert de levier majeur dans la relation entre les enseignants et les étudiants, et a des conséquences éducatives et pédagogiques évidentes, le concept évolue et, d'une fonction administrative, l'évaluation des apprentissages devient le processus privilégié sur lequel l'enseignant s'appuie pour prendre des décisions pédagogiques dans le quotidien de sa salle de classe. C'est pourquoi on s'intéresse beaucoup, depuis quelques années, à connaître les pratiques des professeurs en évaluation des apprentissages en salle de classe. On veut connaître et comprendre les compétences et les habiletés des professeurs en évaluation et on veut décrire ce qu'ils font lorsqu'ils évaluent les apprentissages des étudiants.

Plusieurs signes témoignent d'une convergence d'intérêt sur les pratiques des professeurs en évaluation des apprentissages. Aux États-Unis, les activités et les publications autour du concept de "classroom assessment" se multiplient de façon impressionnante. Lors de son colloque de 1992, le National Council on Measurement in Education (NCME) a tenu un très grand nombre d'ateliers sur ce même sujet, dans la foulée de la publication, en 1990, des "Standards for Teacher

Competence in Educational Assessment of Students" (American Federation of Teachers, 1990). En 1991 et 1992, le NCME a fait une étude panaméricaine des compétences des professeurs en mesure et évaluation des apprentissages et en publiera les résultats à l'automne 1993. Depuis 1992, cette même association analyse les compétences des administrateurs scolaires en matière d'évaluation des apprentissages afin de mettre au point des standards de compétence adaptés à ce groupe d'intervenants.

Plus près de nous, l'Association pour le développement de la mesure et de l'évaluation en éducation (ADMEE) a fait son colloque d'octobre 1992 sous le thème des pratiques d'évaluation. De même, l'Association québécoise de pédagogie collégiale (AQPC) fait de l'évaluation le thème de son colloque annuel de 1994.

Structure du rapport de recherche

Le premier chapitre brossera à grands traits les caractéristiques du problème de recherche. Cet exposé sur la problématique nous permettra de dégager les objectifs de notre recherche et de définir le domaine et le type de la recherche ainsi que ses limites méthodologiques.

Le second chapitre décrit l'état de la question. Nous y présentons notre cadre théorique, faisant ressortir les concepts de croyances, d'attitudes et d'altérabilité des croyances et des comportements. C'est dans ce chapitre que nous chercherons à montrer que les croyances sont en amont des pratiques et que tout effort de changement des pratiques doit s'appuyer sur des changements de croyances.

Le chapitre trois présente une discussion du rôle de l'évaluation en éducation. On y montre que ce rôle évolue à mesure que change, dans la société, le schéma global, le paradigme que nous nous donnons pour concevoir l'évaluation en milieu scolaire. Ce chapitre discute les impacts des pratiques d'évaluation, tant sur l'enseignement que sur l'apprentissage, et montre pourquoi il faut se préoccuper de la nature et de la qualité des pratiques d'évaluation en salle de classe.

Le chapitre quatre décrit la méthodologie que nous avons mise en oeuvre pour accomplir cette recherche. On y trouve toutes les caractéristiques de l'échantillonnage, de la cueillette des données et du traitement des résultats.

C'est au cinquième chapitre que nous présentons les résultats de la cueillette des données. La présentation est organisée à partir d'une typologie fondée sur les six champs de compétence que nous avons adoptés et décrits sommairement au chapitre 4 et repris en détail en annexe 7.

Enfin, le sixième chapitre présentera une analyse des résultats fondée sur les données quantitatives et qualitatives recueillies. Nos recommandations sont intégrées à cette analyse.

La conclusion du présent rapport fera une synthèse de la recherche et proposera quelques pistes de recherches ultérieures.

CHAPITRE 1
LE PROBLÈME

1.1 Définition du problème et objet de la recherche

Dans les cégeps, la mesure et l'évaluation des apprentissages des étudiants relève entièrement et exclusivement des enseignants. Ces derniers sont autonomes et les pratiques qu'ils adoptent sont les seules qui prévalent et qui interviennent, tant dans l'apprentissage que dans la sanction des études. Pour beaucoup d'entre eux, cette fonction est source de stress, de fatigue et de préoccupations éthiques. Il est difficile d'enseigner; il est aussi très difficile d'évaluer. La difficulté s'accroît encore lorsque vient le moment d'attribuer la note finale. À cet égard, le rapport Grégoire (Conseil des collèges, 1985), nous rappelle "la tension presque inévitable qui habite l'enseignant au moment de l'évaluation décisive: il ne peut plus s'agir alors d'accompagner l'étudiant, il faut le juger (p. 86)". À cela s'ajoute, aujourd'hui, la crainte de voir l'État s'immiscer dans le domaine de l'évaluation et d'imposer des examens terminaux aux étudiants de cégeps. Les défis auxquels font face les professeurs à l'égard de la mesure et de l'évaluation apparaissent donc, à l'heure actuelle, de plusieurs ordres: technique, philosophique et même politique.

Pour accomplir leurs tâches d'évaluation, les enseignants des cégeps disposent de formations inégales, sinon aléatoires. À l'embauche, on ne requiert pas qu'ils aient une formation en pédagogie. En conséquence, de nombreux professeurs n'ont aucune formation pouvant les guider lors de l'évaluation des apprentissages de leurs étudiants. Ceux qui ont quelque formation l'auront acquise grâce à l'une ou l'autre des "filières" suivantes:

- leur formation universitaire peut inclure un ou des cours en mesure et évaluation;
- le professeur peut s'inscrire, à sa discrétion, à un cours universitaire en mesure et évaluation, à l'université ou dans son cégep, par exemple dans le cadre du programme PERFORMA.

Ainsi, le savoir-faire collectif en évaluation des apprentissages est aléatoire puisque, dans chaque département, le "profil" de formation en

mesure et évaluation est inégal d'un professeur à l'autre: certains ont reçu une formation en évaluation, d'autres pas.

La compétence est aléatoire aussi à cause du caractère facultatif de la participation des professeurs à des activités de mise à jour et de perfectionnement dans leur établissement ou à l'extérieur.

Pour venir en aide aux enseignants confrontés à leurs différents problèmes de mesure et d'évaluation, les conseillers pédagogiques des cégeps proposent différentes activités de formation et de perfectionnement: des entrevues individuelles, des ateliers, des conférences et des activités de perfectionnement de longue durée. Ces activités sont de nature à toucher tant les techniques d'évaluation que les éléments liés à la tâche et à la planification de l'évaluation. Dans chaque établissement, les politiques d'évaluation des apprentissages proposent également des pistes pouvant guider les enseignants à ce chapitre (Commission de l'évaluation, Conseil des collèges, 1990).

Pourtant, il arrive encore régulièrement que des professeurs manifestent des attitudes et des pratiques qui révèlent une méconnaissance des principes les plus élémentaires en mesure et évaluation. C'est cette méconnaissance qui alimente d'ailleurs, en partie, l'argumentation des partisans d'examens nationaux. Si le réseau collégial postule qu'une formation en pédagogie est un préalable facultatif et qu'il peut inciter son personnel enseignant à se donner une formation en cours d'emploi grâce à la mise en place de divers moyens, il semble que ces moyens ne sont pas toujours efficaces et ne permettent pas d'atteindre les objectifs de perfectionnement visés. Cette situation mérite une attention particulière, sachant à quel point l'évaluation des apprentissages joue un rôle-clé en salle de classe. Plusieurs auteurs ont démontré l'immense influence de l'évaluation et des pratiques d'évaluation sur toutes les dimensions du processus d'enseignement et d'apprentissage. (Stiggins, 1992; Crooks, 1988). Par exemple, Applebee (1988) affirme que:

> "(...)tests and their accompanying grades remain one of the most powerful ways in which we teach students what is important in the subjects they are studying (p. 107)."

L'American Federation of Teachers a publié, en 1990, sept énoncés de compétences en matière de mesure et évaluation en prenant appui sur le rôle déterminant que joue l'évaluation des apprentissages:

> "(...) student assessment is an essential part of teaching and good teaching cannot exist without good student assessment (p. 30). "

Si nous sommes conscients de ces enjeux, comment pouvons-nous contribuer à améliorer les pratiques des enseignants à l'égard de la mesure et de l'évaluation et ainsi protéger une liberté pédagogique chèrement acquise?

Fontaine (1988), dans sa thèse de doctorat intitulée "Mesure critériée des croyances des futurs maîtres à l'égard de l'évaluation des apprentissages", pose le problème en affirmant que les apprentissages cognitifs ne suffisent pas à guider correctement les pratiques:

> "La recherche a démontré très clairement que l'échec d'implantation de plusieurs innovations peut être attribué à un manque d'attitudes positives des utilisateurs à l'égard de l'innovation. Les agents de changement auraient investi considérablement dans le développement des habiletés cognitives des utilisateurs sans se préoccuper suffisamment des comportements affectifs appropriés (p. 3)."

Cette observation réfère aux résistances très spécifiques, rencontrées dans les collèges, lors de l'implantation des Politiques institutionnelles d'évaluation des apprentissages.

Fontaine ajoute plus loin:

> "(...) les comportements des professeurs face à l'évaluation des apprentissages ne peuvent être modifiés uniquement par l'énoncé d'une politique, par la disponibilité d'une instrumentation ou même par la formation à l'utilisation de cette instrumentation. Il faut se préoccuper de la formation des croyances et des attitudes en ce domaine (p. 5)."

Pour comprendre et modifier les pratiques des enseignants en matière de mesure et d'évaluation, Fontaine (1988) affirme, à partir des théories conceptuelles portant sur les attitudes, les croyances et les valeurs, qu'il faut s'intéresser d'abord aux croyances des enseignants si l'on veut amener quelque changement dans les pratiques.

Les croyances semblent donc jouer un rôle capital dans l'orientation des enseignants et paraissent incontournables dès qu'on envisage d'intervenir sur les pratiques. Nous savons qu'aucune recherche, à ce jour, n'a fait l'inventaire détaillé des croyances et des pratiques des enseignants des cégeps relativement à l'évaluation des apprentissages des étudiants. Par conséquent, nous avons choisi, dans un premier temps, d'identifier les croyances et les pratiques des enseignants à l'égard de la mesure et de l'évaluation des apprentissages pour, ensuite, tenter de dégager des pistes de perfectionnement dans ce domaine. Nous croyons que cette démarche permettra d'amorcer une prise de conscience collective de la situation chez les enseignants et d'aider les divers intervenants à planifier des actions de formation plus efficaces qui prendront en compte le rôle des croyances dans la détermination des pratiques de mesure et d'évaluation. Ces actions de formation pourraient ainsi amener chaque enseignant à remettre en question ou à confirmer ses croyances et ses pratiques.

1.2 Les objectifs de la recherche

Voici les objectifs de recherche que nous avons définis et qui, nous l'espérons, nous permettront de mieux comprendre les différents paramètres de l'évaluation des apprentissages au collégial et, notamment, de guider la planification d'activités de perfectionnement adaptées aux besoins des enseignants et enseignantes des cégeps.

1. Identifier les croyances des enseignants des cégeps à l'égard de l'évaluation des apprentissages en salle de classe.

2. Décrire les pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages et ce, en situation régulière de salle de classe.

3. Dégager des pistes de perfectionnement qui prennent en compte les croyances et les pratiques des enseignants du collégial en matière d'évaluation des apprentisages en salle de classe.

1.3 Le type de recherche privilégié

La recherche est de nature descriptive. Le chapitre 4, traitant de la méthodologie, montre que nous utilisons deux méthodes de recueil de données: le questionnaire à choix multiples et l'entrevue de groupe. Le questionnaire est distribué à un échantillon aléatoire stratifié composé de 1,690 enseignants de 47 collèges et campus québécois. Il vise à obtenir des informations sur les croyances et les pratiques des enseignants à l'égard de l'évaluation des apprentissages et sur leur intérêt face à certaines activités de perfectionnement dans ce domaine. Il comporte des énoncés de croyances, des questions sur les pratiques en évaluation et des propositions de thèmes de perfectionnement.

Des entrevues de groupe font suite à une première analyse des résultats et sont menées auprès de groupes de professeurs constitués en groupes

réacteurs. Ceux-ci sont invités à réagir aux résultats de la recherche, à les commenter et à les valider à partir de leur expérience pratique.

1.4 Les limites de la recherche

1.4.1 Limites du domaine

De nombreuses variables, tant objectives que perceptuelles peuvent intervenir pour influencer, voire déterminer les pratiques des professeurs de cégeps en évaluation des apprentissages. Parmi celles-là, la lourdeur de la tâche, tant objectivement mesurable que subjectivement perçue, est maintes fois citée dans les débats sur la condition enseignante et sur la qualité de l'enseignement (Robitaille et Maheu, 1993). Le statut de l'enseignant et le caractère précaire de nombreux postes, l'image de soi et la conception de l'enseignement comptent aussi parmi les déterminants des pratiques. La formation et le sentiment de compétence interviennent aussi.

Nous choisissons de ne pas étudier plusieurs de ces variables intervenantes, soit parce qu'elles sont déjà bien documentées, soit parce qu'elles sont hors du champ de nos objectifs qui, rappelons-le, sont focalisés sur la formation et le perfectionnement en mesure et évaluation.

Dans cet esprit, nous savons que la correction des examens constitue une tâche lourde et fastidieuse et la qualité de certaines études récentes (Conseil des collèges, 1985; Conseil supérieur de l'éducation, 1991; Ranjard, 1984) nous dispensent de le démontrer encore une fois.

Plutôt, nous choisissons d'étudier le problème des pratiques en mesure et évaluation dans les limites de la question de la formation et du perfectionnement en mesure et évaluation car c'est à ce champ d'intervention que nous croyons pouvoir contribuer utilement comme conseillers pédagogiques.

La description des croyances et des pratiques à l'égard de l'évaluation des apprentissages en salle de classe ne saurait être exhaustive car nous avons

choisi de n'aborder que certains aspects de l'évaluation. En effet, suite à nos nombreuses lectures de périodiques, de monographies et de rapports de recherche, nous avons constaté que certains champs de croyances et de pratiques avaient été particulièrement étudiés et décrits: tout ce qui concerne les motifs pour lesquels les enseignants évaluent, les outils d'évaluation et la production de la note. En fait, nous avons réalisé que les recherches portaient beaucoup plus souvent sur le pourquoi de l'évaluation et sur les "produits" de l'évaluation que sur le "processus" de l'évaluation. Nous avons alors décidé, tout en maintenant l'étude de toutes les composantes de l'évaluation, d'accorder une place privilégiée aux différents aspects liés au processus de l'évaluation et aux variables pouvant toucher la qualité de l'évaluation, par exemple, la fabrication des outils d'évaluation, la détermination des critères d'évaluation et de correction, les "contaminants" de l'évaluation. De même, nous avons choisi de limiter la portée de notre étude aux activités d'évaluation ayant cours sur le campus des collèges. Ici, l'évaluation en milieu de stage est exclue, car ce milieu impose des contraintes différentes de celles de la salle de classe. À cet égard, nous pensons que la problématique de l'évaluation en milieu de stage devrait faire l'objet d'une étude spécifique et cela fera l'objet d'une de nos recommandations en conclusion du présent rapport.

1.4.2 Limites méthodologiques

Une première limite, inhérente à la méthodologie du questionnaire, fait en sorte que les répondants n'ont pas toute la latitude souhaitée pour nuancer leurs réponses puisque l'essentiel du questionnaire propose des choix de réponses et des échelles ordinales.

Ensuite, pour favoriser une certaine cohérence dans le profil de réponses des nombreuses questions du questionnaire, nous demandons aux répondants de choisir, dans l'ensemble de leur tâche, un cours témoin et de répondre à tout le questionnaire en fonction de ce seul cours témoin. Le choix que feront les professeurs en amènera vraisemblablement plusieurs

à avoir l'impression, à la fin du questionnaire, de n'avoir pas "tout dit" et de n'avoir pas l'opportunité de témoigner d'autres réalités appartenant à un autre groupe d'étudiants ou à un autre type de cours enseigné.

Une troisième limite nous vient du principe de la désirabilité sociale. Le fait d'avoir à décrire ses croyances ou ses pratiques, même dans un questionnaire anonyme et à choix de réponses, peut amener des répondants à adoucir certains aspects de la réalité: des "souvent" peuvent devenir des "parfois"; des "accords" peuvent nous parvenir sous forme de "ne sais pas". Cette limite, reconnue dans la méthodologie des questionnaires devrait mener le lecteur à comprendre que les résultats de cette recherche sont fondés sur ce que les professeurs veulent bien nous dire.

Le sujet lui-même de notre recherche peut constituer une quatrième limite qui s'ajoute aux trois précédentes. Certains enseignants sollicités ne répondront pas à notre questionnaire à cause du sujet traité. D'aucuns auront une attitude si négative à l'égard de l'évaluation des apprentissages ou du concept d'évaluation qu'ils ne voudront pas y collaborer. D'autres auront peur de l'usage que nous ferons des données et choisiront de ne répondre à aucune de nos questions. Les uns afficheront une peur personnelle fondée sur l'hypothèse qu'on pourrait les identifier, et les autres participeront à une peur plus collective en prétendant que les résultats de la recherche pourraient servir de prétexte à des décisions administratives touchant l'ensemble des professeurs.

Quant aux entrevues de groupes réacteurs, elles demeurent soumises à la subjectivité des interlocuteurs. Les enseignants interrogés seront des volontaires qui nous donneront leur interprétation, leur vision des résultats au questionnaire et qui tenteront de mettre en relation les croyances et les pratiques. Nous espérons que cet exercice permettra une certaine cohérence quant à l'interprétation des données et fera émerger des pistes de perfectionnement intéressantes et pertinentes.

1.5 Définition des termes

Dans notre questionnaire, nous faisons usage de plusieurs concepts et expressions qui appartiennent au domaine de l'évaluation des apprentissages en milieu scolaire. De ce fait, nous avons postulé que ces concepts étaient connus des professeurs des cégeps. Sachant que ces derniers n'ont pas nécessairement reçu une formation formelle en matière de mesure et évaluation, nous ne pouvons qu'espérer que ces mêmes concepts sont aussi compris. Mais nous savons aussi que nous soumettons notre questionnaire à un groupe hétérogène de professeurs et, pour camper minimalement les concepts les plus essentiels, nous y avons ajouté, dans la section intitulée "Comment remplir ce questionnaire", les trois définitions suivantes:

- l'**évaluation** se réfère à toute activité qui permet à l'enseignant de porter un jugement sur les apprentissages des étudiants;
- l'**évaluation formative** s'intéresse à la progression de l'étudiant, pendant la formation, afin d'ajuster l'enseignement ou de corriger les difficultés d'apprentissage;
- l'**évaluation sommative** mesure la somme des apprentissages en vue d'une décision de promotion.

Ces définitions sont extrêmement sommaires car elles correspondent à un compromis. D'une part, nous ne voulons pas "donner les réponses" avant de poser les questions alors que, d'autre part, nous sentons le besoin de proposer des dénominateurs communs à l'ensemble des répondants. Il s'agit donc, pour nous, de maintenir un équilibre précaire mais fonctionnel, entre trop en dire et ne pas assez en dire.

Le processus de validation du questionnaire nous permet de nous assurer que les concepts utilisés dans l'ensemble du questionnaire sont raisonnablement "clairs" et "pertinents". À cet effet, nous référons le lecteur au chapitre de la méthodologie, ci-après.

Mais à cause précisément de ce phénomène par lequel nous soumettons un questionnaire à des personnes, en postulant qu'elles pouvaient y répondre en toute connaissance de cause, nous analyserons les données recueillies en s'appuyant sur ce même postulat. C'est pourquoi il nous semblerait oiseux de proposer, dans ce rapport de recherche, un glossaire des termes utilisés. Nous nous retrouverions dans la situation du professeur qui fournit, le jour même de l'émission des bulletins, le corrigé de l'examen sommatif de fin de session.

Nous choisissons donc de référer le lecteur à l'abondante et excellente documentation disponible au Québec pour y trouver les définitions les plus admises en matière de mesure et évaluation. Tout particulièrement, les auteurs suivants constituent des sources fiables et satisfaisantes:

- Legendre (1993)
- Morissette (1993)
- De Ketele (1991)
- Scallon (1988)

La bibliographie de ce rapport donne les coordonnées précises de ces sources.

CHAPITRE 2
ÉTAT DE LA QUESTION

Depuis quelques années, plusieurs chercheurs ont proposé des théories visant à expliquer de quelle façon s'élaborent les pratiques d'évaluation en salle de classe. Ces écrits tendent à démontrer que les croyances et les attitudes comptent parmi les principaux déterminants des pratiques d'évaluation et que ce sont les croyances qui sont en amont des attitudes et, de là, des pratiques. Le présent chapitre explore ces écrits afin de décrire tant les pratiques des professeurs que leurs croyances en mesure et évaluation des apprentissages.

Nous explorons d'abord le concept de croyances et son rapport avec le concept d'attitude pour ensuite montrer que les pratiques d'évaluation reposent fortement sur des structures de croyances. Dans la perspective de nos objectifs de recherche, le rapport entre les pratiques d'évaluation et les croyances sous-jacentes pose donc la question de l'altérabilité des croyances avant celle de l'altérabilité des pratiques.

2.1 La question des croyances

2.1.1 Le concept de croyances

De nombreux auteurs ont cherché à clarifier le concept de croyances, notamment en essayant de le distinguer du concept de savoir (knowledge). Selon Frank Pajares (1992), certains ont associé la croyance à un jugement évaluatif, à un affect. D'autres en font un construit cognitif.

Déjà, au niveau du vocabulaire, on a parlé des croyances des gens en utilisant de nombreuses expressions. Parmi celles-là, on trouve: opinions, conceptions, attitudes, théories personnelles, théories implicites, dispositions, préconceptions, idéologies, axiomes, etc. C'est pourquoi Pajares dit du concept de croyances que c'est un concept confus et encombré.

Pajares (1992) remonte aux écrits de John Dewey pour donner une première distinction entre "croire" et "savoir". Selon Dewey, une croyance est une réponse qu'on se donne devant l'inconnu. Cette réponse, on la crée lorsqu'on ne dispose pas d'un savoir objectif

(knowledge) pour répondre à une question. Si un professeur n'a pas de formation en pédagogie ou s'il rencontre un problème ou une situation pour lequel sa formation ne lui donne pas de réponse, ce professeur va improviser une solution à ce problème en créant une réponse personnelle à son problème ou en adoptant une réponse toute faite si son entourage lui en propose une qui lui soit convenable.

2.1.2 La relation entre les attitudes, les croyances et les pratiques

Dans le cadre de sa thèse de doctorat portant sur une mesure des croyances des futurs maîtres à l'égard de l'évaluation des apprentissages, Fontaine (1988) a fait une excellente analyse comparative des théories et des recherches sur les croyances et les attitudes. De ce fait, elle a dégagé les pistes permettant d'explorer les relations qui peuvent exister entre les croyances et les pratiques. Cette démarche entreprise par Fontaine (1988) soutient notre position. Comme nous l'avons expliqué au chapitre 1, nous pensons que les croyances sont en amont des attitudes et des pratiques et qu'il est nécessaire, pour modifier ou enrichir les pratiques des enseignants, de travailler d'abord au chapitre des croyances.

D'autres auteurs présentent, certes, des recherches portant sur les croyances ou les attitudes à l'égard de l'évaluation mais ils n'expliquent pas leur préférence pour l'étude des croyances ou des attitudes, ni l'intérêt qu'il y a d'étudier l'un ou l'autre de ces aspects du domaine de l'affectif. À notre connaissance, Fontaine est la seule à avoir fait un choix clair à cet égard et à l'avoir justifié (Fontaine, 1988; Fontaine et Trahan, 1990).

C'est à partir des travaux de Bim et de Rokeach que Fontaine (1988) a défini les concepts de croyances, d'attitudes et de valeurs. La croyance

> " (...) réfère à la perception, par l'individu, d'une relation quelconque entre deux choses ou entre une chose et une caractéristique de celle-ci. La chose peut-être qualifiée de bonne ou mauvaise, vraie ou fausse, désirable ou indésirable, etc. (p. 13)"

L'attitude est "(...) une organisation relativement durable de plusieurs croyances quant à un objet ou à une situation spécifique, prédisposant l'individu à agir d'une certaine manière privilégiée (p. 17)".

Fontaine complète en expliquant qu'une valeur "(...) est une croyance durable à l'effet qu'un mode de conduite spécifique ou une fin de l'existence spécifique est personnellement ou socialement préférable à son opposé ou à sa réciproque (p. 24)."

L'attitude est, en quelque sorte, une synthèse des croyances qui la composent (au moins deux croyances) alors que la valeur est constituée d'une croyance unique d'un genre particulier. Une valeur est une croyance mais une croyance n'est pas nécessairement une valeur. La croyance est donc l'élément de base des attitudes et des valeurs. Elle est crédible et, puisqu'elle donne du sens à la réalité, elle sert à fonder la pensée et l'action (Richardson, 1991). Il est donc logique que tout travail de formation et de perfectionnement soit fondé sur les croyances et sur les liens qu'il y a entre celles-ci et les pratiques.

2.1.3 Genèse des croyances

Dans leur recherche sur les pratiques en évaluation formative, Rose Burelle et autres (1991) ont étudié les théories sur les prédispositions à agir, afin de faire une analyse des déterminants des pratiques en évaluation formative. S'appuyant sur la théorie de l'action raisonnée d'Ajzen et Fishbein (1980), ces auteurs contribuent à clarifier la structure des "événements" qui expliquent le comportement d'un individu dans une situation donnée.

La théorie de l'action raisonnée est fondée sur le postulat que, de façon générale, les humains sont des êtres raisonnables faisant systématiquement usage de l'information dont ils disposent. Même si les humains agissent souvent sous le contrôle de l'habitude, leurs comportements sont généralement précédés d'une réflexion sur les implications possibles de leurs actes sociaux. Pour cette raison, il est possible de prévoir et d'expliquer les comportements d'une personne si les facteurs qui déterminent ce comportement sont connus. Ces facteurs sont endogènes et correspondent, en gros, à la séquence suivante:

- le comportement d'une personne est déterminé par son intention de comportement;
- l'intention de comportement est déterminée par les attitudes du sujet à l'endroit du comportement et par ses normes subjectives;
- ses attitudes sont formées de croyances et ces croyances portent sur les conséquences du comportement et sur les évaluations de ces conséquences. Dans la théorie de l'action raisonnée, le mot "croyance" s'apparente au terme "opinion";
- les normes subjectives sont formées à partir des croyances relatives aux attentes de ses principaux référents (les personnes ou groupes importants à ses yeux) et de sa motivation à s'y conformer.

En décrivant le modèle de Ajzen ef Fishbein, qui relie les comportements à des attitudes et des attitudes à des croyances, Burelle et autres (1991) montrent que divers types de croyances peuvent être générées, en fonction de leur source:

- des croyances peuvent être issues de l'expérience directe. Ce sont des croyances de nature descriptive, proches de la certitude: "la neige est toujours blanche";

- des croyances sont plutôt inférentielles lorsqu'elles sont fondées sur des relations logiques ou sur des relations constatées antérieurement. Si un enfant pleure, on peut croire que c'est parce qu'il est triste;

- d'autres croyances, enfin, sont dites informationnelles. On peut adopter une information reçue d'une tierce personne, un intermédiaire, et la faire sienne, en faire une croyance personnelle. Cette croyance informationnelle est dite comportementale si elle porte sur un comportement: "faire de l'évaluation formative oblige l'enseignant à faire plus de corrections". Par contre, on la qualifiera de normative si elle porte sur une norme, comme: "il faut pratiquer l'évaluation formative".

2.1.4 Altérabilité des croyances

Selon Burelle et autres (1991), cette théorie de l'action raisonnée s'inscrit dans un vaste ensemble de théories de l'apprentissage. Ces théories partent toutes du même postulat de l'altérabilité des attitudes, mais elles se distinguent les unes des autres au plan des moyens proposés pour obtenir des changements.

Selon la théorie de l'action raisonnée, on considère l'attitude comme une disposition affective, un affect, qui émane d'une organisation particulière de croyances. Plusieurs théories d'apprentissage et plusieurs techniques de communication persuasive postulent l'altérabilité des attitudes en plaçant les croyances au centre de la cible des actions de l'intervenant ou du formateur. Cela pose donc la question de l'altérabilité des croyances. C'est là qu'on perçoit la limite du modèle d'Aizen et Fishbein (1980). Ceux-ci se sont limités à essayer de décrire la chaîne d'événements qui expliquent les liens entre les comportements et les croyances.

Autant Fontaine (1988) que Burelle (1991) affirment que c'est la théorie de Rokeach qui peut contribuer le mieux à une réflexion sur l'altérabilité des

croyances et des attitudes. La théorie de Rokeach permet de formuler le corollaire suivant: puisque les croyances sont à l'origine de tout comportement intentionnel, toute influence exercée sur les croyances déterminantes d'un comportement donné est susceptible de se répercuter sur ce comportement, soit pour le susciter, soit pour le modifier.

Pour autant qu'on puisse identifier et mesurer les croyances, on peut supposer qu'une intervention sur la formation des croyances pourrait modifier les comportements des professeurs dans le domaine de l'évaluation en salle de classe.

La théorie de Rokeach est particulièrement utile, ici, parce qu'elle propose une compréhension analytique des croyances. Selon Fontaine (1988), Rokeach introduisait, en 1980, son analyse des types de croyances en posant trois hypothèses:

1. Les croyances ne sont pas toutes également importantes pour un individu; elles varient en importance selon leur position par rapport à un centre.

2. Plus une croyance est centrale, plus elle sera résistante au changement.

3. Plus une croyance qui change est centrale, plus ce changement se répercute sur l'ensemble du système de croyances.

La force d'une croyance peut être inférée en fonction de ses connexions avec d'autres croyances et ces jeux de connexions détermineront l'importance du rôle spécifique de chaque croyance dans l'adoption d'une attitude. Plus une croyance donnée est fonctionnellement reliée à d'autres croyances, plus elle a d'implications sur les autres croyances, plus elle sera dite centrale. Il y a quatre catégories de connexions, pouvant décrire les degrés de "centralité" des croyances.

1. Les croyances qui concernent directement l'existence propre d'un individu et son identité dans le monde physique et social sont dites existentielles. Les croyances existentielles sont présumées avoir plus de liens fonctionnels avec d'autres croyances et avoir des conséquences plus importantes que les croyances non existentielles.

2. Les croyances existentielles peuvent êtres partagées ou non partagées avec d'autres personnes. Les croyances existentielles partagées sont présumées avoir plus de liens fonctionnels avec d'autres croyances et avoir des conséquences plus importantes, que les croyances existentielles non partagées.

3. Plusieurs croyances sont apprises, non par rencontre directe avec l'objet de croyance mais indirectement par l'intermédiaire de personnes ou de groupes agissant comme référents. Ces croyances sont dites dérivées. Les croyances dérivées sont présumées avoir moins de liens fonctionnels avec d'autres croyances et avoir des conséquences moins importantes que les croyances dont elles sont dérivées.

4. Plusieurs croyances portent sur des objets choisis plus ou moins arbitrairement par goût et sont perçues comme telles par ceux qui les professent. Les croyances se rapportant aux matières de goût sont présumées avoir moins de liens fonctionnels avec les autres croyances et avoir des conséquences moins importantes que les croyances ne se rapportant pas aux matières de goût.

Pajares (1992) résume ce modèle de façon très simple: les humains ont diverses croyances de diverses intensités et l'importance de chacune est fonction de ses liens avec les autres ("Human beings have differing beliefs of differing intensity and complex connexions that determine their importance" p. 318).

Dans cette recherche, les croyances relatives aux pratiques d'évaluation en salle de classe relèvent de la troisième catégorie, c'est-à-dire les croyances dérivées. Ce sont précisément ces croyances que Stiggins (1989)

qualifie de folkloriques. Fontaine (1988) réfère aussi à ces croyances dérivées lorsqu'elle affirme que les pratiques d'évaluation et de notation sont, dans la réalité, beaucoup plus "ancrées" dans la tradition que dans la formation. C'est pourquoi, ajoute-t-elle, la formation académique ou les énoncés de politique concernant l'évaluation et la notation sont souvent, à eux seuls, improductifs car ils ne prennent pas en compte les croyances qui soutiennent ces pratiques (p. 5).

2.1.5 Comprendre les attitudes pour guider le changement: Le modèle de Pryor

Dans une étude récente sur des stratégies de mesure des attitudes, Pryor (1992) cherche à décomposer le mécanisme de la formation des attitudes. Il commence par rappeler une définition admise du concept d'attitude, à savoir que c'est une prédisposition envers un objet ou un comportement. Cette prédisposition est apprise.

À partir d'un cadre théorique proposé par Fishbein dans les années soixante, Pryor propose un cadre théorique fondé sur la définition suivante de l'attitude: une attitude est l'affect positif ou négatif à l'égard d'un aspect visible de l'environnement. Selon lui, cette théorie peut donc servir à étudier soit l'attitude à l'endroit d'un objet (par exemple, les mathématiques), soit l'attitude à l'égard d'un comportement spécifique (par exemple, s'inscrire à un cours de mathématiques). Cette théorie peut servir à expliquer la formation des attitudes, autant à l'endroit d'un objet que d'un comportement. L'attitude à l'endroit d'un objet est formée d'un jeu de croyances à propos d'attributs de l'objet ainsi que d'une évaluation de chaque attribut. De même, l'attitude à l'endroit d'un comportement est formée d'un jeu de croyances à propos des résultats attendus de ce comportement ainsi que d'une évaluation de chacun de ces résultats.

Ces couples croyances-évaluations sont additionnés et peuvent constituer une mesure de l'attitude à l'égard de l'objet ou du comportement. Fondé sur ce principe du jumelage de croyances et d'évaluations, Pryor (1992) a

développé une stratégie de mesure d'attitudes qui l'a amené à affirmer que ça peut prendre jusqu'à neuf (9) croyances pour mesurer une attitude.

Le modèle retenu par Pryor confirme le rôle fondamental des croyances dans la formation des attitudes. De plus, puisqu'il permet d'illustrer une à une les croyances qui contribuent à la formation d'une attitude, ce modèle peut être extrêmement utile dans toute stratégie de changement. Il peut éclairer, guider le choix des informations à fournir pour mener à un changement d'attitude dans un sens donné.

2.1.6 Genèse des attitudes et cycle d'évaluation

On peut faire un parallèle entre le cadre théorique de Pryor (1992) sur la formation des attitudes et le modèle en trois temps du cycle de l'évaluation.

Nous avons illustré que, dans le cycle de l'évaluation en éducation, il y a trois temps: la mesure, l'évaluation et la décision (Howe, 1991).

> **Premier temps:** la mesure, qui consiste à recueillir des informations et à les rendre signifiantes, habituellement au moyen de symboles (chiffres, lettres, etc.)
>
> **Deuxième temps:** l'évaluation, le jugement de valeur, fondé sur une comparaison entre les données recueillies et un critère.
>
> **Troisième temps:** la décision, l'action.

Dans son cadre théorique, Pryor (1992) définit aussi trois temps dans la formation d'une attitude: la croyance, l'évaluation et l'attitude elle-même. Nous pensons que la structure du cadre théorique de Pryor équivaut au cycle de l'évaluation en éducation. Puisqu'elle est un construit cognitif, une croyance équivaut à une information, comme celle qu'on obtient lorsqu'on mesure une habileté. L'évaluation des attributs d'un objet ou l'évaluation des résultats attendus d'une action correspond au jugement de valeur que l'on fait sur des informations reçues. La

prédisposition à agir pourrait équivaloir au stade de l'action. Cela suggère le raisonnement suivant:

- si on définit qu'une croyance peut équivaloir à une information qui est fixée à propos de quelque chose, disons un préjugé;
- si on définit que l'attitude est une prédisposition à agir, donc que l'attitude équivaudrait au niveau "décision" ou "action";
- si on définit le 2e temps, l'évaluation qu'on fait d'un attribut ou du résultat d'une action, comme étant un jugement de valeur résultant d'une comparaison entre la croyance et un critère quelconque;
- alors, les trois temps de la formation d'une attitude pourraient être vus comme étant du même modèle que les trois temps du processus (cycle) d'évaluation.

	Le cycle de l'évaluation en éducation	La genèse des attitudes
Premier temps	mesure	croyance
Deuxième temps	évaluation (jugement de valeur fondé sur la comparaison entre l'information et un critère)	évaluation (jugement de valeur fondé sur la comparaison entre la croyance et un critère)
Troisième temps	décision (ou disposition à agir)	attitude (ou prédisposition à agir)

Mais il y a un maillon de la chaîne qui reste à confirmer pour que le parallélisme soit parfait. Quel est ce critère auquel on comparerait la croyance pour qu'on parle vraiment, ici, d'une évaluation? Si Pryor a raison, on a une croyance à propos d'un attribut de quelque chose (ex: les mathématiques, c'est difficile, ou le chocolat c'est bon pour la santé) et on comparera cette croyance à un critère quelconque pour décider, par

exemple, d'avoir telle ou telle attitude à l'égard de ce quelque chose. Ce critère de comparaison peut-il être une valeur? Si les valeurs jouent un rôle dans la genèse des attitudes, ce rôle peut-il être un rôle de déclencheur, de critère d'évaluation, permettant de porter un jugement de valeur et, de là, permettant l'adoption d'une attitude particulière?

2.1.7 Persistance des croyances

Selon Pajares (1992, p. 317), plus les croyances d'une personne sont anciennes, plus cette personne a tendance à mettre en oeuvre des stratégies cognitives qui protègent ces croyances. Cette personne agit pour que ses croyances survivent et ce, même si on lui prouve que ses croyances sont fausses. Par exemple, on entendra des choses comme: "mon professeur faisait comme cela et je m'en suis très bien tiré".

Les croyances se perpétuent d'elles-mêmes et ont tendance à persévérer malgré les éventuelles contradictions qui viennent de l'expérience, de la raison ou de l'éducation. Lorsqu'elles datent de plusieurs années, les croyances deviennent difficiles à altérer.

Si une croyance a été acquise de façon folklorique, dans le département ou n'importe où ailleurs dans l'exercice de la profession, cette croyance cherchera à survivre et résistera à des efforts de formation en cours d'emploi. Cela pose la question de savoir si l'on peut faire en sorte que les professeurs acquièrent, dès le début de l'exercice de la profession, une formation scientifique solide afin d'éviter que, en l'absence de connaissances fondées sur la science, ils n'adoptent des croyances éventuellement fausses et, avec le temps, presqu'inchangeables. Cet argument met en cause le principe de la formation en pédagogie chez les professeurs de cégeps. Si les professeurs entretiennent, entre eux, de fausses croyances, il est probable que ces fausses croyances auront la vie dure et résisteront au changement.

Une croyance, c'est en quelque sorte un succédané à la vérité. On la prend pour une vérité. Et si elle a servi de base à nos actions et à nos

attitudes pendant longtemps et avec satisfaction, si elle est parvenue à répondre à nos questions, à nous orienter et à nous stabiliser, il nous sera difficile d'accepter qu'on nous dise qu'il faudrait en changer.

On sait que les gens agissent à partir d'un ensemble complexe de croyances et de connaissances objectives. On sait aussi qu'on peut faire évoluer les connaissances de quelqu'un par un grand nombre de moyens: la formation académique, les livres, etc. Mais Pajares (1992, p. 311) rappelle que les croyances ne changent pas nécessairement à mesure que les connaissances de quelqu'un changent. Elles ont leur propre vie, elles obéissent à leurs propres lois. Si, dans le système complexe de croyances et de connaissances qui sous-tendent une attitude ou une pratique, l'un des ingrédients déterminants était une croyance, cette croyance pourrait ne pas changer et elle pourrait donc empêcher qu'un changement s'opère dans l'attitude ou la pratique visée. Elle pourrait, à elle seule, entraîner le statu-quo, comme si elle avait un droit de véto.

2.1.8 Profession enseignante et altérabilité des croyances

Dans toutes les professions, des croyances se perpétuent au fil des années. Lorsqu'une personne entre dans une profession, son allégeance aux croyances qui règnent dans la profession sera plus ou moins prononcée, selon que la personne aura eu, auparavant, des contacts plus ou moins forts avec la profession. L'allégeance à des croyances préalables est plus faible chez les avocats, chez les pharmaciens ou chez les architectes, par exemple, parce que l'environnement de travail et la culture propre de ces professions leur sont nouveaux. Ils devront s'apprivoiser aux lieux où s'exercent ces professions et s'approprier, lentement, les connaissances et même les croyances qui y circulent.

Mais la situation en enseignement est fort différente. Les nouveaux enseignants connaissent la salle de classe depuis l'école primaire. Ils y ont développé des préconceptions sur ce que c'est que d'enseigner. Tout nouveau professeur est déjà un initié et est déjà émotivement engagé dans le giron de croyances anciennes. C'est pourquoi l'allégeance aux

croyances préalables est forte chez les professeurs (Pajares, 1992, p. 323) et c'est pourquoi l'altérabilité des croyances, chez les professeurs, est si difficile.

2.1.9 Les croyances et les pratiques des enseignants en évaluation des apprentissages

Les cadres théoriques des auteurs évoqués ci-haut aident à comprendre les paramètres dont il faut tenir compte dans toute entreprise visant à proposer des pistes de perfectionnement aux enseignants des cégeps.

Il est évident que les croyances des enseignants en matière d'évaluation des apprentissages ont un impact sur leurs pratiques. Fontaine (1988) et Pajares (1992) démontrent que la croyance est l'élément de base des attitudes et des valeurs. Il est donc logique que le travail de formation se situe au niveau des croyances et des liens entre celles-ci et les pratiques. Puisque les croyances sont à l'origine de tout comportement intentionnel, toute influence exercée sur les croyances déterminantes d'un comportement donné est susceptible de se répercuter sur ce comportement, soit pour le susciter, soit pour le modifier. Les travaux de Pryor (1992), évoqués plus haut, confirment le rôle-clé des croyances tant dans l'explication de la formation des attitudes que dans la planification des changements d'attitudes.

2.2 La question des pratiques

2.2.1 Apprendre à évaluer

Lorsqu'on pense à l'acquisition d'une compétence dans un domaine quelconque, on conçoit deux grandes catégories de sources de formation:

- la formation académique;
- l'apprentissage expérientiel.

Il en va de même pour les compétences professionnelles des enseignants et, à l'égard des compétences en évaluation des apprentissages, plusieurs auteurs ont décrit et caractérisé le cheminement qu'ont pris les enseignants pour acquérir des compétences en mesure et évaluation.

Formation académique

La formation académique des professeurs de cégeps en matière d'évaluation des apprentissages n'est pas documentée. Mais quelques recherches, à l'extérieur du Québec, donnent une idée de la popularité du sujet.

Dans ses sessions de formation en "classroom assessment", auprès de groupes de professeurs, Stiggins (1991) observe que la moitié, sinon les trois quarts des professeurs inscrits, disent avoir complété au moins un cours universitaire en évaluation des apprentissages. Mais presque tous ajoutent que cette formation n'était pas adaptée à leurs besoins réels et n'était pas très utile dans le cadre de la salle de classe.

Dans l'ensemble, aux États-Unis, la formation en mesure et évaluation n'est pas offerte à tous les enseignants. Des enquêtes montrent que la moitié seulement des programmes de formation des maîtres contiennent un cours obligatoire en évaluation des apprentissages (Boothroyd, 1992; Schafer, 1991; Schafer et Lissitz, 1987). Ces études rapportent que les professeurs qui ont suivi ces cours en mesure et évaluation considèrent que les notions apprises sont rarement celles qui seraient importantes du point de vue des besoins réels en situation de salle de classe, de laboratoire, de stages. D'autres études nous apprennent que, lorsque les cours en mesure et évaluation sont facultatifs, les professeurs ont tendance à les éviter (Marso et Pigge, 1992; Jett et Schafer, 1992).

Dans le réseau collégial québécois, on ne sait pas combien de professeurs ont suivi, à l'université, au moins un cours en mesure et évaluation. Nous savons par ailleurs que la formation en psychopédagogie est facultative, lors de l'embauche, et que les professeurs en exercice sont encouragés à se donner du perfectionnement en pédagogie (Conseil supérieur de l'éducation, novembre 1991, pages 39-46). Le programme PERFORMA de l'Université de Sherbrooke est la formule de perfectionnement en cours d'emploi la mieux connue dans le réseau collégial, mais sa fréquentation n'est pas obligatoire. Depuis 1975, jusqu'en avril 1991, près de 600 professeurs du réseau collégial ont cumulé les 30 crédits du certificat de perfectionnement en enseignement collégial (Ducharme et Forcier, 1991). Cela correspond à 5% des 12,000 professeurs actuels du réseau collégial. À ces nombres, s'ajoutent les très nombreux autres professeurs qui suivent des cours du programme PERFORMA, mais qui n'ont pas accumulé les 30 crédits du CPEC. Puisque les profils de formation sont individualisés et sont modelés, entre autres, par l'offre de cours dans chaque collège, certains de ces professeurs auront suivi, en cours d'emploi, un ou des cours en mesure et évaluation. D'autres, pas.

Apprentissage expérientiel et folklore

Selon Wilk et Carey (1983), repris par Fontaine (1988, p. 4), les pratiques d'évaluation et de notation sont ancrées dans la tradition plutôt que dans la formation et les programmes de formation des maîtres accordent peu d'attention à cet aspect.

Beaucoup de professeurs, en effet, ont acquis par apprentissage expérientiel, tant comme professeurs que comme étudiants, la plus grande part de leur savoir-faire en évaluation des apprentissages. Selon Anderson (1989), les techniques apprises informellement constituent la base principale des pratiques évaluatives des collèges. Ce sont des connaissances de nature folklorique car elles se transmettent entre professeurs, de génération en génération (Stiggins, 1989) et ne coïncident pas nécessairement avec les pratiques recommandées par les experts en sciences de

l'éducation ou en mesure et évaluation. Ces savoir-faire, les professeurs d'aujourd'hui les ont observés alors qu'ils étaient étudiants. En cours d'emploi, ils ont aussi emprunté à leurs collègues plusieurs techniques de mesure et d'évaluation. Ce folklore est perpétué dans les structures d'enracinement, généralement dans les départements, et constitue des traditions qui règnent dans le milieu de l'enseignement (Pajares, 1992, p. 324).

2.2.2 Les pratiques en évaluation des apprentissages

Autonomie dans l'évaluation

De nombreuses recherches ont été conduites, depuis plusieurs années, afin de connaître et de comprendre les pratiques des enseignants en matière d'évaluation des apprentissages en salle de classe. Toutes ces recherches ont pour but de favoriser une compréhension des besoins de formation et de perfectionnement.

Plusieurs recherches ont démontré qu'il y a une corrélation positive entre la compétence des professeurs en évaluation et le fait qu'ils aient suivi au moins un cours de formation en mesure et évaluation. Par exemple, Boothroyd, McMorris et Pruzek (1992) estiment qu'approximativement, 50% des professeurs ont suivi un cours en évaluation des apprentissages, que 53% se disent capables d'identifier des items défectueux dans un examen, et que 44% seulement manifestent des compétences qui correspondent à des standards de qualité en évaluation des apprentissages.

Ces démonstrations de compétences mitigées en mesure et évaluation inquiètent de nombreux auteurs (Boothroyd, et autres, 1992; Schafer, 1991; Stiggins et autres, 1989; Stiggins 1991) car il est démontré que les professeurs ont besoin de générer et d'interpréter de l'information dans le cadre de leur enseignement. En effet, les professeurs prennent, quotidiennement, de nombreuses décisions et celles-ci seront d'autant plus valides qu'elles seront fondées sur des évaluations de qualité. Parmi les

très nombreuses tâches qui exigent des décisions fréquentes, citons, entre autres:

- l'écriture et les mises à jour des plans de cours;
- le suivi et l'encadrement des étudiants;
- le choix et l'amélioration du matériel didactique;
- l'évaluation de l'atteinte des objectifs d'apprentissage;
- l'autoévaluation de la planification et de la prestation de l'enseignement dans une perspective d'amélioration continue.

Plusieurs chercheurs notent que, la plupart du temps, les professeurs fondent leurs décisions sur des informations très informelles et dont ils ne peuvent pas juger la pertinence, la validité, la fiabilité (Airasian, 1991).

Dans l'ensemble de la scolarité d'un étudiant, les tests ou examens sont extrêmement présents et ont une influence énorme sur tout le processus d'enseignement et d'apprentissage. Une étude canadienne (Anderson, 1989) estime qu'un étudiant de Colombie-Britannique est soumis, en moyenne, à près de six cents (600) examens entre le primaire et le second cycle universitaire. De tous ces examens, une vingtaine seulement auront été rédigés par des Comités d'experts: ce sont les examens centralisés dits "du Ministère", au niveau secondaire. Ainsi, la grande majorité des évaluations que subissent les étudiants sont préparés par les enseignants, individuellement, et visent l'évaluation d'apprentissages spécifiques.

Nous savons que les enseignants tiennent à faire eux-mêmes leurs instruments de mesure, refusant dans certains cas, d'entendre parler de toute forme externe d'évaluation (examens centralisés, tests "du Ministère", par exemple). Cette attitude est présente au Québec. Le rapport Grégoire (Conseil des collèges, 1985) en fait état et conclut: "C'est donc l'enseignant qui, à toutes fins pratiques, est le sélectionneur du type d'évaluation" (p. 84).

Et les professeurs valorisent beaucoup les évaluations "maison". Une enquête française (Bigard, 1980) montre que les enseignants tiennent aux examens: "Ils prétendent en effet que sans les examens - c'est-à-dire sans eux, enseignants - non seulement les étudiants ne sauraient rien, mais ne sauraient pas qu'ils ne savent pas" (p. 369).

Instrumentation

Les professeurs utilisent plusieurs méthodes pour évaluer les apprentissages et rédigent, en conséquence, plusieurs types d'instruments de mesure. Déjà, le rapport Grégoire sur la pratique professionnelle des enseignants des cégeps (Conseil des collèges, 1985) démontre que plusieurs modes d'évaluation sont utilisés. Sans aller jusqu'à analyser les motifs des choix des professeurs, le rapport Grégoire constate que:

- les processus "traditionnels" sont dominants: les examens à questions ouvertes et les textes à écrire sont les plus fréquents;
- les tests objectifs sont beaucoup moins utilisés, quoique rencontrés deux fois plus souvent au secteur professionnel qu'au secteur général;
- les enseignants utilisent plusieurs sources de données pour évaluer les apprentissages et transmettre une note finale.

Ce dernier constat trouve un écho dans la recherche d'Anderson (1989) qui, en Colombie-Britannique, étudiait les liens entre les principes recommandés en mesure et évaluation, d'une part, et les pratiques en salle de classe, d'autre part.

En rappelant une étude ontarienne de 1976 qui aurait conclu que les professeurs du niveau secondaire de cette province cherchent une variété d'instruments et de procédures pour évaluer les étudiants, Anderson fait la même observation en précisant toutefois que l'instrument de mesure le plus utilisé, selon son étude, est le test objectif "maison".

Cette étude d'Anderson est faite auprès d'une population différente de celle du rapport Grégoire. Anderson a questionné des professeurs de sciences de 4ième, 7ième et 10ième années sur leurs attitudes et pratiques en mesure et évaluation en classe. Comme dans les cégeps, ces professeurs utilisent plusieurs types d'instruments:

- dossiers anecdotiques;
- tests à questions à développement;
- rapports de laboratoire;
- projets de recherche;
- travaux en bibliothèque.

Selon une étude de Ménard (1990), les étudiants de cégeps répondent en général, sauf en éducation physique, à 3 ou 4 examens par session. Dans ces examens, ce sont les professeurs de philosophie, de physique et de français qui font surtout usage de questions à développement long, alors que les professeurs d'éducation physique et de soins infirmiers semblent être les plus grands usagers de questions à choix multiples.

Dans le cadre d'évaluations qui font appel à des dispositifs d'observation avec grilles d'appréciation, des recherches ont montré que des procédures importantes n'étaient pas mises en oeuvre dans au moins le tiers des cas observés. Les étudiants ne sont pas informés à l'avance des critères de performance, les procédures et les critères de correction ne sont pas définis à l'avance. Dans plus de la moitié des cas, les jugements sont fondés sur une seule observation et, dans 40% des cas, les professeurs se fient à leur mémoire pour appuyer des décisions fondées sur des anecdotes (Stiggins et Bridgeford, 1985). Les auteurs précisent que ces observations proviennent de réponses fournies par les professeurs eux-mêmes et sont donc possiblement entachées du principe de la désirabilité sociale, ce qui permet de croire que les proportions de pratiques incorrectes sont probablement sous-estimées.

Notation

Plusieurs études ont démontré que les professeurs ne sont pas objectifs lorsqu'ils calculent la note finale d'étudiants à risque d'échec. Pour ce type d'étudiants, ils semblent avoir tendance à être plutôt indulgents et cherchent des moyens de donner une "chance au coureur". Le Conseil des collèges (1985) a étudié de façon particulière les politiques personnelles des enseignants à l'égard de l'évaluation. On y constate que les enseignants sont nombreux à chercher un équilibre entre l'indulgence et l'intransigeance et qu'ils optent, en général, pour une attitude indulgente:

- 76% donnent une chance au coureur;
- 70% s'entretiennent ultimement avec le malheureux qui échoue;
- 34% hésitent à inscrire définitivement un échec (p. 86).

Cette tendance à l'indulgence est aussi décrite dans l'étude de 1985 de Christine Barre-de-Miniac (p. 82) qui, là-dessus, va un peu plus loin que le rapport Grégoire en cherchant à expliquer les hésitations à accorder un échec. Les enseignants auraient évoqué la baisse du niveau de connaissances des élèves, la transformation des critères, la volonté de ne pas décourager les élèves et, même, de les valoriser.

Si ces constats sont faits en France, ils le sont aussi dans les collèges du Québec. Les observations du rapport Grégoire (Conseil des collèges, 1985) confirment nettement que nos enseignants sont sensibles, particulièrement au moment de corriger des travaux ou examens, aux acquis tant culturels qu'académiques des étudiants. Certains disent qu'ils sont moins bien préparés qu'avant et qu'il ne faut pas les pénaliser indûment pour ces lacunes. D'autres diront que de mauvaises notes peuvent les décourager et qu'au contraire, il faut dans certains cas, que la note participe à les encourager à persévérer.

L'attitude d'indulgence se manifeste parfois jusque dans un processus peu usuel au collégial: la normalisation des notes. Le Conseil des collèges, dans son rapport sur l'harmonisation du secondaire et du collégial (1989), rappelle une observation du rapport Grégoire, à l'effet que, s'il n'y a pas de processus de "normalisation" des notes au collégial, "cela ne veut pas dire que le professeur (...) n'ajuste pas ses critères d'évaluation au moment de la correction" (p. 34).

De plus, l'adoption de pratiques qui bénéficient d'une "forme ou l'autre de souplesse" (rapport Grégoire, p. 86) pourrait expliquer, au moins en partie, certaines réticences à définir les objectifs des programmes en termes de compétences attendues. Si des professeurs tiennent à conserver toute liberté d'indulgence ou de rigueur pour pouvoir, au besoin, "donner une chance au coureur" (Conseil des collèges, 1985, p. 8), ils peuvent percevoir que des énoncés de compétences attendues constituent des contraintes, des limites à leur marge de manoeuvre: "nous sommes souvent demeurés pudiques dans l'évocation de la ou des compétences et davantage dans leur utilisation comme critère de mesure des performances des étudiants" (Morin, 1989, p. 34).

Dans une recherche auprès d'un échantillon des 16,000 enseignants du secondaire, dans l'état de Virginie, Frary et autres (1992) démontrent que l'immense majorité des enseignants évaluent les apprentissages dans une perspective normative et que les notes contribuent à placer les étudiants en rangs, des plus "forts" aux plus "faibles". Dans cette recherche, la grande majorité est favorable à l'utilisation, pour l'attribution de la note finale, de composantes étrangères aux objectifs d'apprentissage: les progrès réalisés pendant l'année scolaire, l'effort consenti. De même, près du tiers des enseignants sont d'accord à ce que la qualité de la participation en classe puisse intervenir dans la note finale.

Dans une seconde recherche, cette fois auprès des enseignants de niveau collégial, au Virginia Polytechnic Institute, Cross et Frary (1992) observent les pratiques suivantes:

1. Lorsqu'ils ont à décider du sort d'un étudiant en "zone grise", entre une note de passage et une note d'échec,

   a. 75% des professeurs interrogés ont indiqué qu'ils ont tendance à augmenter la note si des efforts exceptionnels méritent d'être récompensés;

   b. 50% des professeurs ont tendance à donner la note inférieure à un étudiant dont le comportement est jugé inacceptable;

   c. chez 30% des répondants, une attitude négative envers le cours influencera à la baisse la note finale attribuée;

2. Près de 50% des professeurs de ce collège américain attribuent la note finale en prenant en considération la qualité de la participation en classe.

Nava et Loyd (1992) démontrent que, lorsqu'ils corrigent des examens ou des travaux et attribuent des notes, les professeurs sont biaisés en fonction du sexe des étudiants. En effet, les professeurs incluent dans la note finale des variables externes à l'apprentissage, par exemple l'effort, le comportement en classe ou en équipe, la présence. Mais ces variables interviennent de façon différenciée, selon que l'étudiant est un garçon ou une fille. En effet, selon la recherche de ces deux enseignantes, les professeurs ont tendance à être:

- plus sévères envers un étudiant fort (capable) mais qui n'a pas fait d'effort, surtout si c'est une fille;
- plus indulgents envers un étudiant moyen qui a fait des progrès, surtout si c'est un garçon;
- plus indulgents envers un étudiant faible qui a fait des progrès, surtout si c'est une fille.

Cette utilisation différenciée de critères extérieurs à l'atteinte des objectifs est jugée incohérente par ces chercheurs qui considèrent que cela

constitue un problème important dans le processus de notation. Ces pratiques incohérentes et différenciées constituent un biais qui contamine et pollue la note finale et qui alimente la perception, chez certains, que l'évaluation est inéquitable.

Les notes sont comme les mots: elles servent à communiquer un message. Les notes doivent avoir un sens et doivent être comprises tant par les professeurs que par les étudiants. Elles doivent signifier la même chose pour tout le monde; elles doivent avoir une signification univoque si l'on veut qu'elles communiquent effectivement ce qu'elles doivent communiquer. Puisque ce sont les professeurs qui décident des éléments constitutifs d'une note, ceux-ci doivent indiquer clairement et à l'avance les critères qu'ils choisissent.

Cette utilisation de variables étrangères à l'atteinte d'objectifs d'apprentissages lors de l'attribution des notes et les implications évidentes sur l'équité et la représentativité des notes sont aussi décrites par Brookhart (1991), Chase et Wakefield (1984), ainsi que par Agnew (1985). Ce dernier a observé que c'est dans un milieu économiquement défavorisé et fréquenté particulièrement par des enfants de minorités ethniques que l'on a le plus recours à des variables externes à l'apprentissage pour attribuer les notes finales: le comportement en classe, la participation, la présence, l'effort. Dans ce milieu, le poids de ces variables est plus élevé que le poids des critères d'atteinte des objectifs. La situation est inversée dans les milieux plus favorisés: les professeurs y utilisent plutôt l'évaluation de l'atteinte des objectifs pédagogiques pour déterminer les notes finales.

Crowl et Berkowitz (1985) ont aussi montré que les pratiques des enseignants à l'égard de la notation peuvent varier en fonction de facteurs propres aux étudiants ou relatifs aux enseignants eux-mêmes.

La calligraphie des étudiants ou des fautes langagières dans les textes, par exemple, peuvent moduler les pratiques de correction et de notation des professeurs. De même, ces derniers peuvent adopter des pratiques variables en fonction, par exemple, de leur expérience dans la discipline

enseignée ou de leur besoin d'être approuvé socialement. À cet égard, Crowl (1984) a trouvé une corrélation significative entre le besoin de certains enseignants d'être approuvés socialement et une tendance à donner des notes plus élevées.

La conception du rôle de la notation, la perception des composantes des notes, la prévision des conséquences psychologiques des notes sur les étudiants sont autant de variables qui influencent les choix des professeurs. Selon Rose et Medway (1981), les enseignants qui s'attribuent la part majeure de la responsabilité des réussites et des échecs de leurs étudiants se trouvent à investir davantage dans leur enseignement et ont des étudiants qui réussissent mieux. Ils croient qu'ils peuvent contrôler le comportement de leurs étudiants et adoptent ainsi de meilleures techniques de gestion de classe. Cela se confirme dans l'étude de Crowl et Berkowitz (1985) selon laquelle les conceptions que se font les professeurs de l'enseignement, de l'évaluation et de la notation auront un impact sur l'atmosphère d'une classe ainsi que sur la conduite des étudiants en classe.

2.2.3 Définir les compétences en évaluation

À cause de l'importance cruciale de l'évaluation des apprentissages en éducation ainsi que des incidences des pratiques d'évaluation, tant sur l'enseignement que sur l'apprentissage, plusieurs efforts sont mis en oeuvre, depuis quelques années, pour définir les compétences que devraient avoir les professeurs en matière d'évaluation des apprentissages. Ces travaux sont fondés sur le constat que l'évaluation des apprentissages est une composante essentielle de l'enseignement et qu'un bon enseignement ne peut se faire sans une bonne évaluation des apprentissages des étudiants.

Les standards de compétence

En 1990, un comité tripartite publiait, aux États-Unis, un document décrivant une série de sept standards de compétence touchant les pratiques d'évaluation des apprentissages des étudiants. Ces standards, rédigés conjointement par deux grandes associations nationales d'enseignants ainsi que par le National Council on Measurement in Education, sont publiés, sans restriction de diffusion ou de reproduction, afin de guider les efforts de formation en mesure et évaluation.

Selon ces sept standards, traduits ici informellement, le personnel enseignant doit :

1. Être en mesure de choisir les méthodes d'évaluation les plus appropriées en fonction des décisions pédagogiques à prendre;

2. Mettre au point, développer, les dispositifs de mesure les plus aptes à guider les décisions pédagogiques;

3. Administrer les dispositifs d'évaluation, attribuer des notes et en interpréter les résultats;

4. Utiliser correctement les résultats d'évaluation dans plusieurs contextes:

   - la planification de l'enseignement;
   - la mise à jour des programmes;
   - l'aide pédagogique aux étudiants;
   - la sanction des études.

5. Mettre au point et utiliser correctement des procédures de correction et de notation des évaluations des apprentissages;

6. Communiquer les résultats des évaluations, tant aux étudiants qu'aux autres instances concernées;

7. Reconnaître toute pratique d'évaluation ou toute utilisation des informations issues de l'évaluation qui seraient illégales, inappropriées ou qui violeraient les règles minimales d'éthique.

D'autres auteurs ont proposé des typologies de compétence en évaluation (Schafer, 1991; Stiggins, 1991) ou ont proposé des critères permettant d'adopter des pratiques correctes à l'égard de l'évaluation ou de la notation (Milton, 1986; Popham, 1991; Robertson, 1991; Stiggins et autres, 1989; Terwilliger, 1977; Wilk, 1983; Wood, 1990).

Plus près de nous, des auteurs québécois nous proposent des écrits susceptibles de guider la formation ou l'autoformation des professeurs, des conseillers pédagogiques et des cadres de collèges en matière d'évaluation en salle de classe. Même s'il n'est pas dans notre propos ici de les résumer, ces textes peuvent tous être mis à contribution pour définir ce que devraient être des pratiques correctes et pour aider à résoudre des problèmes de mesure et d'évaluation. Ces auteurs québécois sont, entre autres, Fontaine (1988), Morissette (1984 et 1993), Scallon (1988), Tousignant et Morissette (1990).

Quant à nous, que ce soit pour guider la structuration de la présente recherche ou pour définir le domaine des compétences à étudier dans des formations à offrir aux enseignants, nous avons choisi d'adopter une typologie fondée sur les travaux de Stiggins (1991). Cette typologie est présentée intégralement à l'annexe 7 du présent rapport, et est comparée à celle de Fontaine (1988), de Schafer (1991) et de l'American federation of teachers (1990) dans le chapitre 4 traitant de notre méthodologie de recherche.

2.2.4 Altérabilité des pratiques

Selon Boothroyd, et al. (1992), les responsabilités doivent être partagées à l'égard des compétences des professeurs en matière d'évaluation. Les curriculums de formation doivent, d'abord, être appuyés sur des

orientations et des contenus pertinents et adaptés aux véritables besoins des professeurs en fonction des ordres d'enseignement et des environnements de pratique. Les structures de leadership, associations professionnelles et autorités académiques doivent définir et promouvoir des standards de compétence qui puissent guider la formation et le maintien des savoirs et des savoir-faire.

La dynamique départementale peut, dans certains cas, inhiber le changement, freiner l'évolution des pratiques. Dans notre réseau collégial, le département est, en principe, le lieu privilégié de l'évolution de la pensée et de la pratique pédagogique. Étant le lieu d'enracinement et d'identité collective le plus important dans la profession enseignante au collégial, le département constitue un espace social incontournable dans toute tentative d'évolution des pratiques, incluant l'évolution des pratiques d'évaluation. Mais depuis quelques années, le département serait devenu, selon les chercheurs Martin Robitaille et Louis Maheux (1993), un espace social stérile duquel de nombreux enseignants se dissocient en adoptant des attitudes d'ambivalence ou de retrait.

> "Dans un tel contexte, (...) on note un repli individuel un peu forcé sur l'activité pédagogique proprement dite au point où la tentation serait grande de traiter la pratique enseignante comme une activité de travail strictement individuelle, activité réduite à la seule relation didactique enseignant et enseigné, et imperméable dès lors (...) à des dimensions plus intersubjectives, collectives, sociales" (Robitaille et Maheux, 1993, p. 94).

Cette attitude de repli sur soi témoigne d'une crainte d'être jugé et évalué pas ses pairs:

> "La confrontation avec un milieu départemental peu interactif (...) recèle un avantage stratégique à court terme: celui de légitimer un repli individuel et libre sur la classe, l'univers pédagogique. Les rapports sociaux avec les collègues, le travail collectif comportent certains inconvénients comme le regard des autres sur ses activités et l'obligation de se concerter, à même de multiples rencontres, ce qui procure à autrui un certain contrôle sur ses activités" (p. 96-97).

Les professeurs se préoccupent de la qualité et de la pertinence de leurs pratiques évaluatives et la plupart sont disposés à une évaluation constructive entre eux, pour s'améliorer. Mais les contraintes endogènes de la vie départementale, les tensions internes, le choc des philosophies incompatibles et les priorités divergentes deviennent autant de facteurs d'inhibition du changement dans les pratiques.

La disponibilité de temps suffisant, associée à la lourdeur de la tâche, constitue une autre contrainte, souvent évoquée, sur l'altérabilité des pratiques. Suite à une recherche publiée en 1985, Stiggins et Bridgeford montrent que les trois quarts des enseignants interrogés se préoccupent de la qualité des évaluations qu'ils font en salle de classe mais se disent inquiets du peu de temps dont ils disposent pour effectuer correctement l'évaluation des apprentissages. Paradoxalement, les chercheurs observent que ces professeurs ne posent aucun geste concret pour améliorer leurs méthodes d'évaluation. Les chercheurs concluent que, malgré les préoccupations exprimées, beaucoup de professeurs n'ont pas l'opportunité, le temps, les moyens ou la motivation de poser des gestes d'amélioration de leurs pratiques.

Dans ce chapitre, nous constatons que les pratiques en évaluation des apprentissages ont des impacts importants, tant sur l'enseignement que sur l'apprentissage et qu'elles sont soumises à des biais personnels et interpersonnels tout aussi importants. Nous y apprenons aussi que les pratiques en évaluation des apprentissages sont fondées, d'une part, sur une aléatoire formation professionnelle et, d'autre part, sur un édifice de croyances.

Le présent chapitre explore les écrits sur les croyances et les pratiques en évaluation des apprentissages et démontre que les croyances et les attitudes comptent parmi les principaux déterminants des pratiques d'évaluation et que ce sont les croyances qui sont en amont des attitudes et, de là, des pratiques. Les croyances sont difficiles à corriger mais il est évident qu'elles restent incontournables dans toute entreprise de modification des pratiques.

CHAPITRE 3

ÉVALUATION, ENSEIGNEMENT ET APPRENTISSAGE

Nous avons vu, au chapitre précédent, que les pratiques en évaluation sont fondées sur des croyances et que toute entreprise de réflexion et d'intervention sur les pratiques doit obligatoirement tenir compte des croyances qui leur sont sous-jacentes. Mais pourquoi est-ce si important de se préoccuper de la nature et de la qualité des pratiques d'évaluation en salle de classe? Dans le présent chapitre, nous comptons montrer que l'évaluation des apprentissages a des impacts, des conséquences, tant sur les professeurs et leur enseignement que sur les étudiants et leur apprentissage. L'évaluation en salle de classe est la source d'information qui mène les professeurs et les étudiants à prendre des décisions très variées, relatives à l'enseignement pour les uns, liées à l'apprentissage et au cheminement scolaire pour les autres. C'est pourquoi nous nous préoccupons d'aider les professeurs des cégeps à acquérir ou à maintenir des compétences suffisantes pour que leurs pratiques évaluatives soient utiles, pour eux-mêmes et pour leurs étudiants, et crédibles aux yeux des étudiants, des collègues, des parents et des contribuables.

Pour décrire les influences de l'évaluation des apprentissages, nous devons camper notre décor. La vision de l'évaluation scolaire a changé, emportée par l'évolution du paradigme de l'évaluation en éducation. C'est cette évolution qui nous sert de toile de fond et que nous décrivons ci-après.

3.1 Le paradigme de l'évaluation des apprentissages

3.1.1 Évolution du paradigme de l'évaluation

L'ancien paradigme en évaluation ou le paradigme du "Test Theory"

Dans "Changing Concepts of Educational Evaluation", paru en 1986, Ralph W. Tyler affirme qu'il y a évolution du paradigme de l'évaluation

en éducation. Il décrit, en des contrastes habilement dessinés, les différences entre l'ancien paradigme et le nouveau[1] .

La pratique de la mesure et de l'évaluation en éducation a changé depuis cent ans. Elle a, en fait, suivi les courants qui ont marqué l'évolution de nos sociétés. D'après Tyler (1986), la démocratisation de nos structures sociales, le passage d'une économie rurale à une économie post-industrielle et l'élargissement des savoirs ont eu un impact déterminant sur notre vision du rôle de l'école et de l'évaluation.

Au début du siècle, on affirmait que la plupart des enfants avait un potentiel "d'éducabilité" limité. Des chercheurs, en 1910, avaient produit des tests d'aptitudes visant à prédire la capacité de réussir des enfants admis à l'école. Ces tests nommés "tests d'intelligence" démontraient, pensait-on, que peu d'enfants étaient en mesure de compléter des études de niveau secondaire. L'évaluation servait alors à classer, c'est-à-dire à départager ceux qui étaient éducables de ceux qui ne l'étaient pas. On encourageait les meilleurs à poursuivre leurs études alors qu'on invitait les autres à se diriger vers le monde du travail qui avait d'ailleurs grand besoin de main-d'oeuvre non qualifiée. En 1910, aux États-Unis, seulement 10% des élèves terminaient leurs études secondaires.

Les deux grandes guerres mondiales ont favorisé le développement, voire la prolifération des tests d'intelligence et d'aptitudes. Ces différents

---

1 Un paradigme est défini par Ouellet (1981) comme étant un ensemble d'hypothèses sur lesquelles on organise et développe des théories et des modèles. C'est un schéma global qui sert de repère dans nos investigations, dans nos recherches. Le paradigme est plus général que la théorie ou le modèle. Une théorie ou un modèle est plus spécifique qu'un paradigme. Par exemple, Paul Inchauspé (1992) propose de guider les changements à réaliser dans notre système scolaire en cherchant de nouveaux paradigmes "sur lesquels s'aligneront les acteurs". Ces paradigmes, l'auteur les qualifie d'ensemble de critères de référence. Cette définition implique qu'un ensemble de professionnels dans un champ d'activités partagent implicitement le même cadre conceptuel.

tests servaient au recrutement et au classement des effectifs militaires. En 1945, Hildreth répertoriait l'existence de cinq mille deux cent quatre-vingt-quatorze "Mental Tests" et "Rating Scales". Selon Guba et Lincoln (1991), le domaine de la mesure et de l'évaluation de cette époque se caractérisait comme suit:

1. Les concepts d'évaluation et de mesure se confondaient, car on ne faisait pas d'évaluation sans passer par la mesure, par la psychométrie. Ces concepts étaient considérés comme des équivalents, des termes interchangeables. On parlait même davantage de mesure que d'évaluation. Dans la francophonie, le seul terme de "docimologie" suffisait.

2. Les concepts de mesure et d'évaluation étaient du domaine des sciences exactes. Ils tenaient davantage de la psychométrie que du programme scolaire et du curriculum.

3. La mesure et l'évaluation étaient à la recherche de la différence entre les personnes plutôt qu'à la recherche d'informations susceptibles d'aider à prendre des décisions.

4. La mesure et l'évaluation consistaient en des mesures objectives et standardisées qui servaient de normes de référence. Leur fonction était de placer les individus sur un continum, du plus faible au plus fort. Le score attribué était perçu comme prédisant les possibilités de succès de l'étudiant. Ce paradigme postule que toutes les mesures de performance humaine peuvent être distribuées sur une courbe en forme de cloche, la courbe normale de probabilités.

Mais voilà, le rôle normatif qu'a joué l'évaluation s'avère inadéquat dans le cadre de l'éducation d'aujourd'hui. D'une part, il ne répond pas aux questions et aux besoins des étudiants, des parents et des enseignants qui veulent de l'information réelle sur ce qui a été appris, sur ce qui n'est pas maîtrisé ou sur les progrès réalisés. De plus, il encourage une vision fausse de l'école qui est perçue alors comme un lieu de compétition

plutôt qu'une institution sociale qui a pour fonction d'aider les étudiants à apprendre.

Le nouveau paradigme en évaluation ou l'évaluation intégrée au processus d'enseignement et d'apprentissage

C'est ce même Ralph Tyler, aujourd'hui âgé de 90 ans, qui décrit avec une lucidité et une honnêteté intellectuelle admirables les lacunes de l'ancien paradigme et qui explique comment la recherche et l'évolution de la société nous obligent à tenir compte de nouveaux critères de référence, d'un nouveau paradigme, pour réfléchir à nos pratiques en matière d'évaluation des apprentissages.

Selon Tyler, nos sociétés, autant par principe que par besoin, ont encouragé au fil des ans une plus grande fréquentation scolaire. La démocratisation des structures sociales, l'évolution du marché du travail, l'explosion des savoirs ainsi que la recherche en éducation obligent à redéfinir le mandat de l'école et à en ouvrir les portes à un nombre croissant d'enfants.

D'une part, l'école n'a plus la capacité de transmettre toutes les connaissances. Les problématiques que nous connaîtrons demain n'existent peut-être pas encore aujourd'hui. L'école doit préparer davantage l'individu à apprendre de façon autonome et à résoudre des problèmes complexes dans différentes situations de la vie qu'à cumuler des connaissances factuelles.

D'autre part, la recherche en éducation permet de mieux comprendre le processus de l'apprentissage. Plusieurs études ont démontré qu'il n'y a pas de corrélation entre le quotient intellectuel de l'étudiant et son rendement scolaire. Au cours des années soixante, Bloom affirme que la relation entre ces deux variables est si faible qu'elle peut être considérée comme négligeable et que la corrélation, qui existe entre les tests d'aptitudes et le rendement scolaire, s'explique du fait que les tests d'aptitudes scolaires mesurent les connaissances préalables de l'individu

dans un domaine, connaissances qui sont nécessaires à l'apprentissage, donc liées directement au rendement scolaire. C'est pourquoi les enfants de milieu favorisé obtiennent généralement des résultats meilleurs à ces tests.

Nos sociétés, en se démocratisant, ont facilité l'accès à l'école mais n'ont pas prévu de structures pour faciliter la réussite scolaire et pour répondre aux nouveaux besoins de formation. L'évaluation se devait de changer de paradigme pour être en mesure d'accompagner cette envolée.

Ralph W. Tyler est l'un des premiers à participer à ce changement. Sa principale contribution au monde de l'éducation est d'avoir suggéré d'organiser les curriculums autour d'objectifs de formation. En 1950, dans "Basic Principles of Curriculum and Instruction", il affirme que l'évaluation est essentiellement un processus qui sert à vérifier si les objectifs de formation ont été atteints. Il propose de développer des instruments capables d'évaluer l'atteinte de chacun des objectifs de formation. Ces instruments doivent être objectifs, fiables et valides. Émergent également les concepts de rétroaction et d'évaluation formative. Cette approche constitue une évolution par rapport à l'approche normative en permettant à l'évaluation de jouer un rôle actif dans l'amélioration des programmes de formation. Tyler affirme aussi que les concepts de mesure et d'évaluation sont des concepts distincts, des processus différents et que la mesure n'est qu'une façon de réaliser l'évaluation.

Notre perception de l'évaluation a continué à évoluer, depuis, grâce aux travaux de plusieurs auteurs majeurs, tant en Amérique qu'en Europe (Guba et Lincoln, 1991; De Ketele et al, 1991). Notamment, vers la fin des années soixante, les concepts d'évaluation sommative et d'évaluation formative étaient différenciés, permettant l'émergence d'un modèle d'évaluation centré sur le client. Dans la décennie suivante, on définissait l'évaluation dans un cadre de prise de décision et on démontrait sa contribution au développement des curriculums. Enfin, au cours des années 80, on précisait les trois fonctions de l'évaluation, à savoir: la fonction prédictive (ou diagnostique), la fonction certificative (ou sommative) et la fonction formative (Cardinet, 1986).

Ces contributions permettent de clarifier le rôle de l'évaluation en salle de classe et ses fonctions dans la dynamique de l'apprentissage. Le nouveau paradigme en évaluation est construit autour des nouvelles attentes de la société à l'égard de l'éducation. On croit que toute personne peut apprendre ce que l'école enseigne et que l'ensemble de la population a besoin de ce que l'école peut offrir afin de vivre pleinement sa vie en société. Le nouveau paradigme affirme que l'école peut s'adapter aux besoins des étudiants et que l'évaluation joue un rôle clé dans cette adaptation.

3.1.2 L'évaluation, une aide à l'apprentissage

Si on fait de l'évaluation, quel que soit le contexte, c'est qu'on a des décisions à prendre, c'est qu'on cherche à agir et que, pour ce faire, nous désirons appuyer nos décisions sur des informations valides. En fait, les professeurs prennent constamment des décisions, pour réguler l'apprentissage des étudiants et pour appliquer les correctifs nécessaires dans leur enseignement. Selon Stiggins (1991), un professeur prend une décision, en salle de classe, à toutes les trois minutes en moyenne. Philippe Meirieu observe cette réalité et affirme que:

> "Le pédagogue est (aussi) un décideur qui effectue des choix en vue de réguler de manière efficace l'activité de l'élève. L'évaluation peut alors se mettre au service de l'apprentissage soit directement, en éclairant l'activité de celui qui apprend, soit indirectement, en éclairant les choix de celui dont la mission est de faciliter les apprentissages." (cité par Hadji, 1990)

C'est à ce processus de décision que réfère le concept d'évaluation formative, lui-même au coeur du modèle de la pédagogie de la réussite. Dans une philosophie de pédagogie de la réussite, l'enseignant se pose constamment des questions portant sur l'évolution de son enseignement et, surtout, sur l'évolution de l'apprentissage des étudiants. Dans cette perspective, l'évaluation peut aider à l'apprentissage. Puisque l'évaluation

formative sert à guider des décisions de nature pédagogique (remédiation, enrichissement et régulation du rythme, de la cadence), on peut désormais parler d'apprentissage assisté par l'évaluation (Hadji, 1990).

L'évaluation n'est plus un acte extérieur au processus d'enseignement et d'apprentissage, qui dérange le processus pédagogique du professeur et qui ne sert qu'à couronner le succès ou à confirmer un échec. L'évaluation n'est plus un geste uniquement administratif. Le nouveau paradigme conçoit l'évaluation comme une partie intégrante du processus de l'enseignement et de l'apprentissage. L'évaluation appartient désormais à la pédagogie.

3.1.3 De l'évaluation à l'assessment: Une question d'attitude

Le paradigme de l'évaluation des apprentissages nous amène à nous intéresser à l'attitude qu'un professeur adopte lorsqu'il évalue. D'après Frisbie (1991), si le concept de l'évaluation des apprentissages change, c'est aussi parce que les professeurs désirent qu'il change. Ceux-ci sont conscients de leur responsabilité en regard de la sanction des compétences minimales, ils savent qu'ils posent des gestes qui ont des effets directs sur les perceptions étudiantes à l'égard de l'échec ou du succès, à l'égard de l'étude, à l'égard de ce qui importe ou importe moins dans la matière enseignée ou dans les objectifs éducatifs visés. Dans une étude sur le pouvoir de l'évaluation à contribuer au développement personnel des étudiants, Paris et autres (1991) discutent des qualités que devrait avoir toute évaluation en classe. Ils en proposent trois:

1. L'évaluation devrait être de nature collaborative afin de favoriser l'apprentissage et la motivation. Lorsque les professeurs participent avec les étudiants à l'évaluation en classe, ils peuvent observer leur raisonnement, leur incompréhension, les processus mis en oeuvre au cours d'une résolution de problème. La rétroaction corrective et le renforcement peuvent être fournis tout de suite. Une telle participation est analogue au comportement d'un entraîneur qui dirige, observe, corrige, intervient lorsque son aide est nécessaire, mais se

retire lorsque l'étudiant fonctionne de façon autonome. C'est le concept de "coaching". Dans une telle supervision interactive, l'étudiant prend part aux jugements, aux critères, aux résultats et aux décisions du processus d'évaluation. Et, du même coup, le professeur participe personnellement à l'évolution des compétences.

2. L'évaluation devrait être cumulative. Elle devrait cumuler des données, des informations de manière à ce que l'étudiant puisse dégager un tableau de son évolution, de ses progrès, tout au long de la session.

3. L'évaluation devrait enfin être multidimentionnelle et ne pas porter seulement sur la discipline enseignée, mais aussi sur des habiletés de formation fondamentale, sur des variables affectives et métacognitives. Elle devrait être multidimentionnelle parce qu'il y a interrelation entre l'apprentissage, la motivation et le succès.

En conformité avec l'étymologie latine "d'assessment", les enseignants s'asseoient avec l'étudiant, sont constructifs et aidants. Ils portent des jugements prudents et méticuleux en les fondant sur l'observation d'un ensemble complexe de caractéristiques et de compétences. En classe, un professeur qui "s'asseoit avec" ses étudiants pour leur donner une rétroaction aidante, non pas seulement sur le contenu disciplinaire, mais sur un large spectre d'habiletés et d'apprentissages affectifs, qui collabore et coopère à l'évaluation comme un partenaire, ce professeur fait ce que les anglophones appellent de l'assessment (Forcier, 1991; Alverno College, 1985).

Ce concept est malheureusement intraduisible en français. Cependant, il correspond presque entièrement au concept d'évaluation formative tel que nous le comprenons dans la francophonie. C'est probablement ce qui explique que la documentation anglophone n'utilise plus l'expression

"formative evaluation", fondée sur la nuance entre les fonctions sommative et formative, mais réfère plutôt à l'expression "classroom assessment", fondée sur l'attitude qui préside à ce type d'évaluation. L'évaluation formative, comme l'assessment, se caractérise par l'attitude bienveillante de la personne qui évalue. C'est cette perspective qui permet d'affirmer que l'évaluation formative peut désormais s'inscrire directement dans la didactique et prendre appui sur les interventions pédagogiques, plutôt qu'uniquement sur des instruments formels de mesure (Howe, 1991; Scallon, 1992).

Ainsi, le concept d'évaluation change, non pas à cause des techniques, mais à cause des philosophies et attitudes sous-jacentes. Guba et Lincoln (1991) affirment que tous les impliqués, l'évalué comme l'évaluateur, que ce dernier soit professeur ou administrateur, sont intimement liés dans l'acte d'évaluation. Ils deviennent des "stakeholders". Ils sont parties prenantes.

3.2 Quelques conséquences de l'évaluation

La nouvelle façon de voir l'évaluation des apprentissages en salle de classe nous oblige donc à revoir la conception qu'on se fait de la relation maître-élève. Elle amène aussi à tenir compte, entre autres, des incidences de l'évaluation sur l'enseignement et sur l'apprentissage.

3.2.1 L'évaluation et l'enseignement

L'obligation d'évaluer les apprentissages ainsi que les pratiques mises en oeuvre dans ce processus ont des conséquences sur la tâche des professeurs, notamment aux chapitres de la notation et de la correction, ainsi que sur leurs choix des contenus enseignés et leurs approches pédagogiques.

Les impacts sur la notation

Pour la plupart des professeurs de tous les ordres d'enseignement, l'évaluation des apprentissages constitue l'aspect le plus difficile de leur travail. Selon plusieurs enquêtes, les professeurs craignent que l'évaluation, et particulièrement leur évaluation, n'ait un impact négatif sur l'étudiant, comme personne et comme étudiant, à court ou à long terme (Barnes, 1985; Conseil des collèges, 1985). Plusieurs ont peur de blesser les étudiants.

Dans une excellente étude sur les pratiques de notation, Suzan Brookhart (1991) rappelle qu'une note, attribuée par un professeur, sera d'une part interprétée et, d'autre part utilisée. Or, puisque plusieurs professeurs s'inquiètent de l'utilisation qu'on fera de la note (échecs d'un étudiant pourtant prometteur, abandons d'un profil de formation, difficulté à se trouver un emploi ou à être admis à l'université), certains d'entre eux vont ajouter des variables atténuantes pour dire que l'étudiant a quand même du mérite, a fourni des efforts valables, a du potentiel. Ces autres variables (effort, participation, etc.) ont pour effet direct d'augmenter une note qui serait plus faible si elle était fondée uniquement sur les compétences acquises.

Cela introduit un nouveau problème. Si on s'inquiète des conséquences sociales de l'évaluation, au point d'introduire diverses variables dans la composition d'une note finale, on en arrive à ne plus savoir comment interpréter cette note. Selon Brookhart (1991), plusieurs enseignants sont ambivalents lorsqu'ils se préoccupent du principe de l'interprétabilité de la note, d'une part, et de l'utilisation sociale de cette note, d'autre part. Selon elle, plusieurs auteurs soulignent que ce phénomène remet en question la validité des évaluations et affirment que l'interprétabilité des notes et les conséquences sociales de leur utilisation doivent compter désormais parmi les critères d'analyse de la validité de toute évaluation.

Selon Pamela Moss (1992), il y a émergence, chez les spécialistes en mesure et évaluation, d'un nouveau consensus sur le concept de validité. Selon ce consensus, toute réflexion sur la validité de construit doit,

maintenant, tenir compte de l'interprétation des résultats des évaluations des apprentissages et des conséquences des évaluations en éducation. Invoquant les travaux que Samuel Messick publiait en 1989 sur la validité de construit, Suzan Brookhart (1991) rappelle que la validité d'un instrument de mesure est fonction, entre autres, de:

- la pertinence des interprétations prévues de la note; et
- le degré de pertinence des décisions qui seront prises à partir de cette note.

Par exemple, selon le premier facteur de validité, il faut pouvoir savoir ce que contient la note et ce qu'elle signifie (i.e.: Est-ce une note fondée sur l'effort, la présence en classe, etc?) Selon le second facteur de validité, il faut savoir ce qu'on fait avec cette note. (Veut-on accorder un prix d'excellence? Va-t-on s'en servir pour attester de la compétence terminale de quelqu'un dans un secteur professionnel?)

Aux États-Unis, les professeurs sont de plus en plus conscients de leur responsabilité sociale à cet égard. Plusieurs font l'objet de poursuites judiciaires à cause d'échecs ou de succès injustifiés (De Landsheere, 1988, p. 217-223).

Les impacts sur la tâche

On sait que l'évaluation des apprentissages en salle de classe occupe une portion importante de la tâche des professeurs. Certaines recherches ont tenté d'en mesurer l'ampleur. Selon les contextes, on estime que la mesure et l'évaluation des apprentissages occupe 10% de la tâche des professeurs (Schafer et Lissitz, 1987). Stiggins (1991) va plus loin et affirme que certains professeurs consacrent entre le tiers et la moitié de leur tâche professionnelle à des activités reliées à l'évaluation des apprentissages. En termes de "temps de classe", les activités relatives à l'évaluation des apprentissages occuperaient près du quart du temps alloué aux activités d'enseignement (Jett et Schafer, 1992; Stiggins, 1992, p. 145). Une étude de

Shavelson et Stern (1981) estime que les enseignants prennent des décisions de nature pédagogique au rythme d'une à toutes les deux ou trois minutes et ces décisions sont toutes fondées sur leur propre évaluation de l'apprentissage de leurs étudiants.

Lorsqu'il est question de corrections, tous les textes font état des heures considérables que les enseignants y consacrent. Patrice Ranjard (1984) qualifie de "cauchemar" la correction des travaux et des examens. Dans son étude sur les tâches des enseignants, la correction est classée, par 63% des répondants à son enquête, en tête de liste des tâches particulièrement lourdes.

Le rapport Grégoire (Conseil des collèges, 1985) a fait, et nous croyons que c'est heureux, une distinction entre évaluation et correction, celle-ci n'étant que le processus permettant d'établir une mesure et, de là, permettant de juger, donc d'évaluer. Ceci étant dit, le rapport Grégoire nous donne quelques chiffres aux pages 85 et 86:

- 40% des répondants à l'enquête Grégoire ont investi entre 10 et 25 heures de correction en un (1) mois;
- 37% des répondants ont dû consacrer plus de 25 heures à cette tâche dans la même période;
- on met plus de temps à corriger au secteur général qu'au professionnel;
- les enseignants les plus expérimentés font plus d'heures de correction que les professeurs qui ont moins d'expérience.

Ranjard (1984) parle même d'enseignants qui travaillent plus de 50 heures par semaine à cause des corrections:

> "Rien de plus fastidieux, irritant, desséchant que cette lecture interminable de copies le plus souvent sans intérêt, mal écrites et criblées d'incorrections qui se répètent obstinément malgré toutes les corrections. Car les élèves ne regardent que la note, ne tirant aucun profit des annotations écrites et répètent éternellement les mêmes fautes (p. 68)".

Encore récemment, la correction se voit confirmée en première position au palmarès des tâches désagréables et accaparantes de l'enseignement:

> "Au collégial, l'évaluation, et en particulier la correction des travaux et examens, ressort nettement comme la tâche la plus désagréable (pour 60% des professeurs interrogés). Vient ensuite, mais loin derrière, la participation à des réunions mentionnée par environ un enseignant sur dix." (Conseil supérieur de l'éducation, 1991, p. 38).

Les impacts sur la pédagogie

L'impact des pratiques d'évaluation sur la pédagogie est moins connu mais tout aussi important. Selon Pamela Moss (1992), ce qui n'est pas évalué a tendance à disparaître du curriculum enseigné. Si cette affirmation est fondée au collégial, on peut dire que les examens et les stratégies utilisés pour évaluer ont une influence déterminante sur l'enseignement et l'apprentissage. D'autres auteurs (Crooks, 1988; Gong, 1992), affirment que l'évaluation est un point de levier si important qu'on peut modifier les pratiques d'enseignement simplement en changeant les pratiques d'évaluation.

3.2.2 L'évaluation et l'apprentissage

Guy Romano (1991), ayant réalisé une recherche sur les pratiques d'étude chez nos étudiants de cégeps, note que les étudiants se développent des stratégies d'étude plus ou moins sophistiquées, plus ou moins profondes, selon les examens qu'ils auront à passer en classe. L'examen sommatif,

puisqu'il contribue au calcul de la note finale, a des effets directs sur les stratégies d'étude de la matière. Les choix que font les professeurs en matière de pratiques d'évaluation (instrumentation, fréquence, notation, rétroaction, stratégies, niveaux taxonomiques) déterminent donc, au moins en partie, le type d'étude que font les étudiants.

Les observations de Lundeberg (1991) vont dans le même sens. Si l'étudiant sait que le professeur évaluera ses apprentissages au moyen de questions objectives et qu'il croit que ces types de questions mesurent habituellement des connaissances mémorisées, il aura tendance à étudier plutôt superficiellement. De même, si l'étudiant prévoit que le professeur utilisera des questions ouvertes et qu'il pense que ces types de questions mesurent habituellement des habiletés de compréhension ou d'application, il fera une étude plus en profondeur, plus analytique.

Si l'instrumentation et les stratégies des professeurs "inspirent" le type d'étude des étudiants, les contenus évalués en font autant. Dès la première évaluation, les étudiants apprennent à distinguer ce que le professeur juge important dans la matière et ont tendance à étudier en fonction de cette perception. En effet, beaucoup d'étudiants, peut-être la plupart, vont planifier leur étude en vue de réussir les examens ou toute autre tâche évaluative. Par conséquent, ce qui n'est pas évalué tend à disparaître de l'étude. Les étudiants ont donc tendance à ajuster leurs stratégies d'étude en fonction des stratégies d'évaluation de leurs professeurs (Noizet et Caverni, 1978, 1983). C'est ainsi que beaucoup de professeurs se tirent dans le pied en manquant de cohérence entre ce qu'ils annoncent et ce qu'ils évaluent. Par exemple, les enseignants ont beau insister sur des habiletés de compréhension et d'intégration de leur contenu, ils sabotent leur propre travail d'enseignement lorsque, dans leurs examens, ils demandent aux étudiants de manipuler des données mémorisées (Crooks, 1988, p. 445; McMorris et Boothroyd, 1992).

Toujours selon Crooks (1988), la relation causale entre les pratiques d'évaluation et les pratiques d'étude est si forte que le meilleur moyen de

modifier les comportements d'apprentissage des étudiants est de modifier les pratiques d'évaluation.

Les professeurs savent bien cela et plusieurs tentent d'influencer les pratiques d'étude chez leurs étudiants en adoptant certaines pratiques d'évaluation. Dans une étude récente, Green (1992) démontre que certains professeurs croient que les questions d'examen de type "à développement", lorsqu'elles sont rédigées en vue de mesurer des habiletés intellectuelles supérieures, sont de nature à ne pas encourager l'étude parce qu'elles font appel à la réflexion et à l'analyse. Selon ces professeurs, les étudiants auraient tendance, en quelque sorte, à se fier à leur capacité d'improvisation. En fonction de cette croyance, ces professeurs choisissent de n'utiliser que des examens objectifs mesurant les connaissances de base.

Par contre, d'autres professeurs croient que les élèves étudient plus lorsqu'on utilise des questions à développement et que ces questions se prêtent mieux à la mesure des habiletés cognitives supérieures (Marso et Pigge, 1992). Malgré les contradictions apparentes, ces observations montrent bien que les professeurs cherchent à adopter des pratiques d'évaluation en fonction des effets qu'ils cherchent à produire sur les pratiques d'étude.

3.3 De la docimologie à la mesure et évaluation

Si Ralph Tyler (1986) invoquait des contraintes économiques, sociales et politiques pour "casser" l'ancien paradigme, nous constatons que c'est la recherche en sciences humaines et, particulièrement, l'évolution des sciences de l'éducation et de la psychologie cognitive qui font évoluer notre compréhension du nouveau paradigme. Le passage de l'ancien au nouveau paradigme aura d'ailleurs eu un effet subtil sur notre vocabulaire puisque nous ne parlons plus, aujourd'hui, de docimologie mais de mesure et évaluation en éducation.

Ce glissement de l'ancien paradigme vers le nouveau ne se fait pas non plus sans confusion. Dans le discours pédagogique, on n'est pas toujours capable de distinguer correctement les concepts en évaluation en les rattachant à l'un ou l'autre paradigme. Et pour compliquer les choses, beaucoup ne semblent même pas conscients qu'un paradigme, ancien ou nouveau, préside à ce discours.

CHAPITRE 4
LA MÉTHODOLOGIE

Dans ce chapitre, nous offrons d'abord un aperçu général de la méthodologie employée pour effectuer cette recherche, puis nous expliquons de quelle façon nous avons élaboré et validé notre instrument de cueillette de données. Nous présentons, pour terminer, la composition de notre échantillon ainsi que l'approche privilégiée pour parfaire notre cueillette des données.

4.1 Aperçu général

Cette recherche est essentiellement descriptive. Elle comporte deux volets: un premier s'appuie sur l'utilisation du questionnaire pour effectuer la cueillette des données tandis que le second repose sur la participation de groupes réacteurs comme stratégie de validation des données.

La recherche exploite des données de diverses origines. Selon une nomenclature proposée par Van der Maren (1990, p. 88-89), nos données sont :

- provoquées, c'est-à-dire, obtenues par questionnaire;
- suscitées, c'est-à-dire, sollicitées auprès de groupes réacteurs;
- évoquées, car la documentation disponible est prise en compte.

Le questionnaire s'adresse à un échantillon aléatoire stratifié composé de 1690 enseignants de 47 collèges publics de la province. Il nous permet d'obtenir des informations relatives aux trois objectifs de la recherche.

Les données suscitées proviennent des rencontres avec trois groupes réacteurs composés de plusieurs professeurs de différents cégeps. Ces rencontres, réalisées durant la période d'analyse des données, servent à enrichir l'interprétation des données et à préciser l'orientation de nos recommandations.

Ces deux approches se veulent complémentaires et visent à répondre aux objectifs de recherche suivants:

1. Identifier les croyances des enseignants des cégeps à l'égard de l'évaluation des apprentissages en salle de classe;

2. Décrire les pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages et ce, en situation régulière de salle de classe;

3. Dégager des pistes de perfectionnement qui prennent en compte les croyances et les pratiques des enseignants du collégial en matière d'évaluation des apprentissages en salle de classe.

Les données sont présentées au chapitre suivant. Rappelons que cette recherche n'aborde que les activités d'évaluation se déroulant à l'intérieur des murs des collèges ou, à tout le moins, sur le campus des établissements.

4.2 Élaboration du questionnaire

4.2.1 Extraction des énoncés

Pour élaborer le questionnaire, nous identifions, à l'intérieur des documents et des recherches consultés, un ensemble d'énoncés portant sur les croyances et les pratiques des enseignants à l'égard de l'évaluation des apprentissages. Cette étape est très longue car les Américains et les Canadiens de l'Ontario et de la Colombie-Britannique publient beaucoup sur l'évaluation des apprentissages mais leurs écrits sont difficilement accessibles ou pas disponibles au Québec. Nous devons, à quelques reprises, communiquer directement avec les chercheurs pour obtenir une copie de leur rapport. La multiplication des démarches de ce genre nous retarde considérablement mais nous procédons quand même à une revue exhaustive des écrits.

Suite au dépouillement de cette documentation, nous nous retrouvons avec 278 énoncés de croyances et de pratiques. Nous classons ensuite ces énoncés en deux groupes: les énoncés décrivant des croyances et les énoncés décrivant des pratiques.

4.2.2 Le choix d'une typologie

Pour nous assurer, d'une part, de circonscrire le domaine de l'évaluation des apprentissages et, d'autre part, pour regrouper, distinguer et classer les nombreux énoncés de croyances et de pratiques recueillis, nous avons besoin d'adopter un système de catégories adapté à notre domaine de recherche. La documentation spécialisée suggère plusieurs pistes de catégorisation pouvant s'adapter au contexte des croyances et pratiques en évaluation des apprentissages. Nous choisissons quatre (4) de ces pistes et les étudions au mérite, dans le but d'adopter l'un des systèmes ou encore de nous créer un système bien à nous, inspiré des quatre auteurs étudiés. Les quatre catégorisations étudiées, ici, sont celles proposées par: Fontaine (1988), Schafer (1991), l'American Federation of Teachers (1990) et Stiggins (1991).

Les catégories inspirées de Fontaine (1988):

1. La signification de l'évaluation et la notation;
2. Le rôle de l'évaluation;
3. L'évaluation normative ou critériée;
4. Les effets de l'évaluation: stress, autorité, incitation au travail, compétence, échecs;
5. Les aspects formels et informels de l'évaluation;

6. Les aspects considérés lors de la notation: efforts, stress, indulgence, pondération, normalisation;
7. La valeur des résultats et des notes;
8. L'importance et l'utilité de l'évaluation;
9. Évaluer des cibles différentes: attitudes, compétences, etc.
10. La fréquence de l'évaluation;
11. L'enseignement cognitif;
12. L'enseignement équitable;
13. La responsabilité du professeur à l'égard des notes;
14. Les différences individuelles entre les étudiants;
15. L'autoévaluation;
16. Le rôle d'information des parents.

Les catégories inspirées de Stiggins (1991):

1. L'utilisation de l'évaluation en classe;
2. Les cibles de l'évaluation;
3. Les qualités de l'évaluation;
4. Les instruments d'évaluation;
5. Les dimensions interpersonnelles de l'évaluation;
6. La rétroaction et la notation.

Les catégories inspirées de l'American Federation of Teachers (1990):

1. Le choix des méthodes d'évaluation appropriées aux décisions à prendre en pédagogie;
2. Les habiletés dans le développement des instruments de mesure;
3. L'administration, la correction, la notation, l'interprétation des résultats;
4. Les habiletés à utiliser les résultats dans les décisions pédagogiques: évaluation formative, évaluation sommative, évaluation de programme, planification, etc.
5. Les procédures de notation;
6. La communication de résultats aux intéressés;
7. L'évitement de pratiques illégales, non-éthiques, inappropriées en mesure et lors de l'utilisation de l'information obtenue.

Les catégories inspirées de Schafer (1991):

1. Les concepts de base et la terminologie de l'évaluation: approches, critères de qualité en évaluation, bonne utilisation du vocabulaire;
2. L'utilisation de l'évaluation;
3. La planification et le développement de l'évaluation: définition du domaine, choix des formes appropriées d'instruments de mesure;
4. L'interprétation de l'évaluation: notation, calcul, interprétation, communication, limites de validité;

5. L'utilisation des résultats: calcul de normes et de statistiques;

6. L'analyse et l'amélioration des évaluations;

7. Le feed-back et la notation: utilisation des données à des fins formatives et sommatives;

8. Les conséquences de l'évaluation: éthique et équité.

Voici un tableau comparatif des systèmes de catégories inspirées de ces quatre auteurs, permettant d'analyser l'appariement inter-systèmes, afin de dégager les similitudes et les différences entre ces quatre systèmes.

Tableau 1: Comparaison de quatre catégorisations du domaine de l'évaluation des apprentissages

Fontaine (1988)	Stiggins (1991)	American Federation of teachers	Schafer (1990)
1 et 5		7	8
2 et 8	1	3 et 4	2
9	2	1	2 et 3
	3	2	5 et 6
	4	1	3
4, 6 et 1, 3	5		4
7	6	5 et 6	7

Note: Les nombres renvoient aux énoncés de catégories énumérées ci-haut. Certains énoncés sont invoqués plus d'une fois; d'autres n'ont pas de correspondance chez les autres auteurs.

Après analyse, nous constatons que les 16 catégories de Fontaine conviennent davantage à l'objet de sa recherche, c'est-à-dire, les futurs maîtres sans expérience, qu'au nôtre, les enseignants d'expérience sur le terrain. Dans le système inspiré de Schafer, la première catégorie s'appuie sur des compétences cognitives seulement et est peu aidante pour décrire des pratiques. Cette catégorisation est cependant utile puisqu'elle se réfère probablement aux fondements de certaines croyances en évaluation. Les catégories de l'Americain Federation of Teachers sont, elles, conçues pour correspondre à une séquence d'étapes ou d'événements entourant l'acte d'évaluation des apprentissages. La septième étape concerne la gestion des règles éthiques et nous semble peu pertinente dans le cadre de notre recherche. Les six catégories de Stiggins nous apparaissent les plus intéressantes parce qu'elles sont fondées sur le concept de compétences dans la perspective de l'évaluation en salle de classe, telle que conçue et gérée par le professeur.

Le système de catégories proposé par Stiggins (1991) est retenu, tel quel, parce qu'il:

- est plus axé sur la pratique de l'évaluation en classe que le système proposé par l'American Federation of Teachers qui, lui, inclut des habiletés de jugement;
- est plus axé sur la pratique en salle de classe que le système de Schafer qui inclut des habiletés cognitives;
- comprend toutes les catégories retenues par Fontaine à la fin des études factorielles sur son instrument de mesure, c'est-à-dire, les catégories 4, 6, 7, 8 et 9;

- permet de nuancer les trois questions structurantes que sont le "pourquoi", le "quoi" et le "comment" de l'évaluation:

  Pourquoi évaluer (catégories 1 et 6 de Stiggins):

  1. pour aider à prendre des décisions;
  2. pour aider au processus d'enseignement et d'apprentissage;
  3. pour aider à gérer la classe.

  Quoi évaluer (catégorie 2 de Stiggins):

  les cibles, les objets de l'évaluation menant à la rétroaction ou à la notation.

  Comment évaluer (catégories 3, 4 et 5 de Stiggins):

  la qualité, l'instrumentation, les influences interpersonnelles.

4.2.3 Le système de catégories de Stiggins

Dans le cadre d'une longue étude sur les besoins de perfectionnement des enseignants en matière d'évaluation des apprentissages, Richard J. Stiggins (1991) a fait l'analyse et la description des compétences qui sont nécessaires à tout enseignant qui souhaite: 1) évaluer correctement une gamme variée de résultats d'apprentissage; 2) gérer correctement l'ensemble du processus d'évaluation en salle de classe.

C'est de cette analyse que Stiggins dérive les six champs de compétences que nous choisissons d'adopter comme catégories de classement. Dans le résumé qui suit, nous présentons les six (6) catégories de classement. On trouve, en annexe 7, le texte complet de ces champs de compétences.

1. L'utilisation de l'évaluation en classe:

Cette catégorie se réfère aux buts pour lesquels l'enseignant évalue. Selon les observations de l'auteur, l'enseignant se sert, en général, de l'évaluation pour appuyer ses décisions, pour guider son enseignement, pour gérer la classe, pour contrôler les comportements des étudiants et les faire travailler.

2. Les cibles de l'évaluation:

Il est question, ici, des objets sur lesquels portent l'évaluation. Les cibles, généralement évaluées par les enseignants, sont: la connaissance de la matière, le savoir-faire, les habiletés intellectuelles supérieures et les attitudes. Pour Stiggins, l'enseignant doit savoir comment évaluer chacune de ces cibles.

3. Les qualités de l'évaluation:

Les caractéristiques d'une "bonne" évaluation varient selon le contexte de l'évaluation. Cependant, quelques critères de qualité sont communs à toutes les situations d'évaluation: la correspondance entre l'objet à évaluer et le dispositif de mesure choisi; le contrôle des sources d'erreur de mesure; la concordance entre l'apprentissage visé et les résultats d'évaluation obtenus; l'obtention d'une information dont l'interprétation est claire pour l'enseignant et les étudiants.

Les qualités de l'évaluation ne s'expriment donc pas seulement en termes de fidélité et de validité des instruments de mesure. Les enseignants doivent résoudre des problèmes relatifs à la conception, à l'administration et à la notation des dispositifs de mesure, ainsi qu'à l'interprétation des notes et à l'utilisation qui en sera faite.

4. Les instruments d'évaluation:

Les professeurs utilisent au moins trois formes d'évaluation des apprentissages en salle de classe: les instruments de type "papier-crayon", l'observation et l'échange verbal. Les professeurs doivent savoir se servir de ces instruments tout en comprenant que les règles de validité sont différentes d'une forme à l'autre.

5. Les dimensions interpersonnelles de l'évaluation:

Les variables interpersonnelles ont un impact sur l'évaluation en salle de classe ainsi que sur l'interprétabilité des notes. Certaines contraintes interpersonnelles viennent des étudiants eux-mêmes, d'autres, des enseignants. Par exemple, les attentes et les attitudes envers les individus; la conception de leur rôle dans l'enseignement et dans l'apprentissage, l'évaluation et le contrôle de la salle de classe; le tempérament; la motivation à enseigner et les aptitudes à réagir aux personnes. L'enseignant compétent doit savoir reconnaître et contrôler ces nombreuses dimensions personnelles s'il veut assurer à ses étudiants une évaluation équitable.

6. La rétroaction et la notation:

Les enseignants fournissent continuellement des rétroactions à partir des résultats des évaluations. Les pratiques de notation sont plus appropriées lorsque les professeurs:

- informent à l'avance les étudiants de leur politique et de leurs attitudes à ce sujet;
- construisent la note finale autour des variables pertinentes;
- introduisent dans la note finale des résultats fondés sur des évaluations de qualité;

- maintiennent des données complètes et appropriées quand aux apprentissages de leurs étudiants;
- pondèrent correctement les diverses notes accumulées et établissent convenablement les seuils de réussite afin de traduire en note les compétences atteintes.

Quant aux autres formes de rétroaction (la communication verbale et non-verbale, les annotations sur les travaux écrits et compositions, les grilles d'évaluation sur échelles de jugement), les professeurs ont des pratiques appropriées lorsqu'ils s'assurent que leur rétroaction correspond à ce qui est attendu explicitement dans l'objectif visé, qu'elle est transmise dans le temps approprié et qu'elle est signifiante pour l'étudiant.

4.2.4 Élagage et ajout d'énoncés

Suite au choix d'une typologie, nous décidons de procéder à un premier élagage d'énoncés. Pour ce faire, nous comparons d'abord nos énoncés avec ceux des questionnaires et des documents d'origine. Nous remarquons alors que la plupart d'entre eux portent sur les catégories 1 (l'utilisation de l'évaluation: rôles et fonctions), 4 (les instruments d'évaluation) et une partie de la catégorie 6 (la rétroaction et la notation) de Stiggins. Nous trouvons peu d'écrits traitant des cibles de l'évaluation (catégorie 2 de Stiggins), de la qualité de l'évaluation (catégorie 3 de Stiggins) et de la notation (catégories 5 et 6 de Stiggins). Pour combler ces lacunes, nous rédigeons des énoncés et élaborons de nouvelles questions correspondant aux catégories 2, 3, 5 et 6 de Stiggins.

Nous choisissons donc de mettre l'emphase sur le processus de l'évaluation car il représente le champ des croyances et des pratiques des enseignants le moins connu. Par ailleurs, afin de pouvoir éventuellement faire des liens avec les résultats de recherches antérieures, nous gardons "l'essentiel" des catégories 1, 4 et 6 (production de la note) de

Stiggins. L'essentiel, c'est le plus habituel, le plus typique de la pratique des enseignants du collégial.

4.2.5 Des entrevues de pré-consultation

Des entrevues d'environ une heure, réalisées auprès d'enseignants de notre collège, nous permettent de prendre contact avec l'ordinaire de la pratique et des préoccupations des enseignants. Si la cueillette des données par questionnaire constitue une consultation, ces entrevues représentent, à cette étape, une pré-consultation. Elles permettent de tester les cultures de divers milieux afin d'ajuster, au besoin, le vocabulaire de nos énoncés de questions. Elles nous assurent qu'il y a congruence entre les pistes dégagées de la recension des écrits et la perception qu'ont les enseignants de leurs pratiques (Gadbois et al, 1991). Cette pré-consultation est faite d'octobre à décembre 1991 auprès de douze enseignants d'âge, d'expérience, de statut et de départements différents et ce, à partir d'un questionnaire à questions ouvertes. Nous utilisons la méthode de l'entrevue semi-dirigée (Van der Maren, 1990; Daunais,1986). Le protocole d'entrevues ainsi que le profil des professeurs interviewés sont présentés en annexe 3. Les entrevues sont individuelles et conduites, chacune, par un seul des deux chercheurs. Un enregistrement audio de chaque entrevue est fait, après signature d'un protocole de confidentialité entre les chercheurs et chaque interviewé. Le verbatim de ces entrevues est transcrit et analysé. Ce processus contribue à enrichir le questionnaire en nous permettant de valider, dans le contexte spécifique des cégeps, l'ensemble des énoncés inspirés de la documentation. Cette pratique, qui consiste à obtenir des pistes de questions directement de personnes représentatives de la population visée par la recherche, est confirmée par Pryor (1992):

> "Il est d'une importance capitale d'extraire des énoncés de croyances directement de la population étudiée. Les énoncés suggérés, extraits d'autres sources, la documentation spécialisée ou les experts-conseils, peuvent, probablement, servir à mesurer efficacement les attitudes, mais ce sont des représentants de la population directement étudiée qui pourront mettre en lumière les variables cognitives, culturelles et affectives qui sont sous-jacentes aux attitudes." (traduction libre)

Puis, finalement, nous nous demandons si les informations issues des énoncés de départ et des entrevues peuvent être récupérées au cours d'une activité de perfectionnement. En effet, il nous faut garder en tête que le but de la recherche est de proposer des éclairages susceptibles de guider la planification d'activités de perfectionnement. C'est ainsi que les énoncés, en regard des problèmes de plagiat et de stress aux examens, sont enlevés car ils concernent plus la gestion du comportement étudiant que le processus de l'évaluation. Nous nous posons, à cette étape, les questions suivantes:

1. L'énoncé porte-t-il sur le processus ou le produit de l'évaluation?
2. L'énoncé est-il approprié au contexte de l'évaluation au collégial?
3. L'information découlant de l'énoncé peut-elle avoir un impact sur une activité de perfectionnement?

Parallèlement, nous cherchons à faire en sorte que le questionnaire puisse également être utile au répondant. Nous désirons que ce dernier ait le sentiment d'y apprendre quelque chose et soit stimulé à pousuivre sa réflexion.

Cette première démarche nous permet de réaliser un premier élagage des 278 énoncés de départ et de rédiger le questionnaire destiné aux juges-experts.

4.3 Validation et mise en forme du questionnaire

Le questionnaire est soumis à deux niveaux de validation: une validation du contenu (clarté et pertinence des énoncés) par des juges-experts et une validation d'ensemble, une "répétition générale", par un petit échantillon de la population visée.

4.3.1 Pré-validation

Le questionnaire est pré-validé à la fin du mois de janvier 1992 auprès d'un groupe de consultants-experts que nous recrutons, en novembre et décembre, à partir des compétences respectives de chacun. Ces consultants, au nombre de sept, sont: un conseiller pédagogique de cégep, expérimenté en perfectionnement des professeurs; un membre du Groupe de Recherche-Action de PERFORMA; un expert-conseil en enquête et sondages; un professionnel du Service de la recherche de la DGEC; un membre de la Commission d'évaluation du Conseil des collèges; un conseiller pédagogique d'université, spécialiste en mesure et évaluation; un professeur d'université, spécialiste en mesure et évaluation.

Ces consultants effectuent une lecture critique en regard des aspects suivants: clarté et pertinence de chacun des énoncés de questions; pertinence des catégories. Nous leur demandons, aussi, d'ajouter les commentaires et suggestions qu'ils jugent appropriés. La liste des énoncés est révisée, en tenant compte de leur jugement et de leurs commentaires. L'annexe 2 identifie ces consultants-experts et reproduit les documents examinés.

4.3.2 Mise en page

Nous consultons ensuite un spécialiste en enquête et sondages pour compléter ce processus de validation et pour réaliser une mise en page

professionnelle du questionnaire. Dans cette mise en forme, nous alternons des blocs d'énoncés de croyances avec des blocs d'énoncés de pratiques, de façon à éviter la monotonie de trop longues séries de questions, tout en respectant une économie minimale de consignes de réponses. Dans la constitution de ces blocs, nous cherchons à éviter une symétrie trop évidente entre certains énoncés de croyances et de pratiques, afin d'éviter que la réponse à l'un, n'influence la réponse à l'autre.

4.3.3 Validation de la pénultième version

Nous procédons à l'étape de validation finale du questionnaire en administrant une avant-dernière version à un échantillon de 25 enseignants de notre collège. Cet échantillon a des caractéristiques semblables à celles de notre population échantillonnée. Une dernière correction du questionnaire est réalisée suite à cette étape. Nous respectons avec minutie toutes les étapes de fabrication et de validation d'un questionnnaire, recommandées par les experts dans le domaine (Wolf, 1988). La version finale du questionnaire, telle que distribuée à la population échantillonnée, est présenté en annexe 1.

4.4 L'échantillon

Nous définissons notre échantillon à partir d'une identification de la population réelle des professeurs de cégeps, telle que décrite dans la banque SPOC du Service de l'informatique du M.E.S.S. Le Service des communications de la DGEC met à notre disposition une partie de la version de novembre 1991 de cette banque informatisée, contenant les noms de tous les professeurs de l'enseignement régulier des cégeps. Ces noms sont accompagnés de la discipline enseignée et du cégep d'appartenance.

De là, nous stratifions notre échantillon en nous basant sur les familles des programmes offerts dans le réseau collégial, telles que définies dans

les cahiers de l'enseignement collégial. Au départ, nous avons donc identifié les onze strates échantillonnales suivantes:

101	Sciences biologiques
199	Techniques biologiques
201	Sciences physiques
299	Techniques physiques
301	Sciences humaines (incluant Philosophie)
399	Techniques humaines
401	Administration (pré-universitaire)
499	Administration (technique)
501	Arts (pré-universitaire)
599	Arts (technique)
601	Lettres

En utilisant les méthodes d'échantillonnage les plus admises, nous constituons l'échantillon total en identifiant 1690 professeurs dans la banque SPOC parmi les 11693 qui y sont répertoriés.

Les proportions relatives de professeurs, dans chaque famille des programmes, sont appliquées à l'échantillon total. Ceci nous donne, dans chaque strate, des nombres proportionnels de professeurs échantillonnés.

Tableau 2: Distribution des proportions attribuées à chaque strate

	POPULATION		ÉCHANTILLON		
STRATE	N	%	SOLLICITÉS		RÉPONDANTS
			N	N	taux de réponses
101	979	8.5%	160	66	41.2%
199	1607	13.7%	160	77	48.1%
201	1605	13.7%	160	88	55%
299	1184	10%	160	64	40%
301	2074	17.7%	160	55	34.3%
399	710	6%	160	54	33.7%
401	223	1.9%	130	19	14.6%
499	1064	9%	150	87	58%
501	549	4.6%	150	22	14.6%
599	435	3.7%	140	39	27.8%
601	1263	10.8%	160	64	40%
TOTAUX	11693		1690 * 1510 **	670 ***	

Notes: * professeurs sollicités

** professeurs rejoints (estimé)

*** Il y a eu 670 répondants, mais 35 d'entre eux n'indiquent pas leur discipline d'appartenance (Question 60)

De l'ensemble des questionnaires distribués, 670 nous sont retournés. Puisque nous estimons notre échantillon total à 1510 professeurs, notre taux de réponses est de 44.3%. Ce taux de réponses est supérieur aux taux de 25% à 30% souvent observés dans des sondages par la poste. La liste des cégeps participants ainsi que le taux de réponses obtenues pour chaque cégep se retrouvent à l'annexe 4.

Pendant et après la cueillette des données, notre échantillon total et notre stratification sont modifiés pour les raisons suivantes:

La précision de la banque SPOC. Tous les questionnaires envoyés dans les collèges sont adressés à des personnes dont les noms sont fournis par la banque informatisée SPOC la plus à jour. Dans 19 cégeps, la personne responsable de la distribution des questionnaires nous contacte pour signaler que certains destinataires ne sont pas à l'emploi du collège à ce moment-là et ce, pour l'une ou l'autre des raisons suivantes: congé, mise en disponibilité, décès, transfert de collège. D'un collège à l'autre, les nombres varient, d'un minimum de 3% à un maximum de 22%. En généralisant cette tendance à tous les collèges, nous calculons qu'il y a un écart moyen de 10,6% entre l'échantillon de départ et l'échantillon réellement rejoint. Ainsi, nous pouvons estimer que le taux de précision de la banque SPOC est, au moment de son utilisation, de 89.4%. Nous présumons, en conséquence, que notre échantillon total n'est plus de 1690 mais plutôt de 1510 professeurs. Ce nombre nous assure, avec un intervalle de confiance de 95%, que les données recueillies seront représentatives de la population.

Les strates à réponse faible. Les strates Administration (pré-universitaire) et Arts (pré-universitaire) génèrent un faible taux de réponses, ce qui, lors de l'analyse des données par tableaux de contingence, risque d'invalider plusieurs tableaux. Les règles de validité des tableaux de contingence (Trudel et Antonius, 1991, p. 408-415) nous amènent à décider de regrouper en permanence les deux strates Sciences et Techniques administratives ainsi que les deux strates relatives aux Arts (pré-universitaire et

secteur technique). De onze strates, ces regroupements nous ramènent à neuf strates.

La strate Sciences biologiques. La strate "Sciences biologiques" contient, tel que prévu dans les cahiers de l'enseignement collégial, les disciplines Biologie et Éducation physique. Dès le début des analyses des données, il apparaît que ces deux grandes disciplines affichent des profils de réponses souvent fort différents. Par exemple, l'une utilise "beaucoup" les laboratoires alors que ce n'est pas du tout le cas dans l'autre. Nous choisissons, a posteriori, de séparer "Éducation physique" de sa strate "mère" et d'en faire une strate à part. Cette décision ramène notre échantillon stratifié à une structure en dix strates:

CODES	Appellations complètes des strates
101	Sciences biologiques
109	Éducation physique
199	Techniques biologiques
201	Sciences physiques
299	Techniques physiques
301	Sciences humaines (incluant Philosophie)
399	Techniques humaines
400	Administration (pré-universitaire et technique)
500	Arts (pré-universitaire et technique)
601	Lettres

Il faut souligner que nous avons, également, pondéré l'échantillon de ces 10 strates. En effet, à l'intérieur de chaque strate d'origine, nous notons que les taux de réponses varient, créant a posteriori des écarts entre, d'une part, la représentativité des strates dans la population et, d'autre part, le taux réel de réponses dans les strates. Un facteur de pondération est calculé pour chaque strate afin que l'ensemble des répondants de la strate représentent le poids de la strate dans la population. Ces pondérations, présentées à l'annexe 4, permettent de calculer des fréquences pondérées qui servent, au prochain chapitre, à décrire les données obtenues de l'ensemble des répondants.

Toutes les analyses stratifiées qui sont présentées, dans ce rapport, sont fondées sur cette structure en dix strates.

4.5 La distribution des questionnaires

Pour des raisons d'économie budgétaire, nous mettons en place un réseau de distribution qui prend appui sur la collaboration d'un répondant dans chaque établissement. Nous avons contacté les conseillers pédagogiques ou adjoints pédagogiques des 47 collèges participants, pour leur demander de distribuer, par le courrier interne de leur collège, la lettre de pré-avis, le questionnaire ainsi que le carton de rappel.

Grâce à ce réseau de distribution, nous expédions, fin mars 1992, une lettre de pré-avis aux enseignants qui font partie de notre échantillon provincial. Cette lettre les informe qu'ils sont choisis au hasard pour répondre à un questionnaire portant sur l'évaluation des apprentissages des étudiants (voir le questionnaire en annexe 1). Nous distribuons ensuite notre questionnaire au cours du mois d'avril 1992. Les cartes de rappel sont postées trois semaines après l'envoi des questionnaires afin de relancer les non-répondants. Les cartes de rappel ont eu un effet positif, augmentant les retours de questionnaires à un total de 670. Notons que ces cartes de rappel sont envoyées à tout notre échantillon

car, les questionnaires étant anonymes, nous ne pouvons identifier et cibler les non-répondants.

4.6 Les groupes réacteurs

Selon notre planification originelle, nous devons effectuer des entrevues individuelles, après une première étude des données quantitatives, afin de valider ou d'interpréter les données. Cependant, les données quantitatives s'avèrent fort considérables: la base de données est faite de 193 variables. L'analyse des données exige un travail beaucoup plus long que prévu.

Nous choisissons donc de constituer des groupes réacteurs, et de recueillir des données qualitatives comme dans des groupes de discussion (Geoffrion, 1992). Cette stratégie de faire appel à des groupes réacteurs constitue un "retour à la source" (Guba et Lincoln, 1991, p. 110; Huberman et Miles; 1991, p. 442), et a pour but de tester la crédibilité, la plausibilité de nos données et ce, auprès des informateurs eux-mêmes. Les groupes réacteurs sont des outils qui nous permettent de confirmer les résultats du questionnaire.

Nous convoquons trois groupes réacteurs, l'un dans un grand cégep de la région de Montréal, le second dans un grand cégep de la région de Québec et le troisième dans un cégep de taille moyenne, en province.

Par le truchement du conseiller pédagogique de chaque établissement, les professeurs de ces cégeps sont invités à participer à ces groupes réacteurs en sachant que nous allons leur demander de réagir à des données recueillies par questionnaire afin de les interpréter, les valider, les illustrer d'exemples.

Les trois groupes réacteurs sont rencontrés sur une période étalée entre janvier et avril 1993. La participation est, en moyenne, de 25 professeurs dans chaque groupe. Nous savons que, d'un groupe à l'autre, les questions ou commentaires des professeurs vont nous amener à questionner

de nouveaux aspects des données quantitatives. Ainsi, par le phénomène d'émergence décrit par Guba et Lincoln (1991) dans le cadre de la recherche naturaliste, chaque groupe réacteur constitue, pour les chercheurs, une source de données relativement imprévisible. Le protocole de rencontre de ces trois groupes est, dans sa structure fondamentale, constant: présentation commentée des données quantitatives; échange et discussion avec les participants. Cependant, la dynamique interne des groupes réacteurs permet l'expression de profils de questions et de commentaires qui varient d'un groupe à l'autre. Avec l'assentiment de chacun de ces groupes, les échanges sont enregistrés sur bande audio et les données qualitatives obtenues sont intégrées aux analyses présentées au chapitre 6.

CHAPITRE 5
LES RÉSULTATS DE LA RECHERCHE

5.1 Composition du chapitre

5.1.1 Contenu du chapitre

Nous décrivons ici les données recueillies à partir du questionnaire vu au chapitre précédent. Les données sont présentées dans leur ensemble puis par strates à l'intérieur de tableaux de fréquences. Cet inventaire des données permet l'atteinte des deux premiers objectifs de la recherche, à savoir:

1. Identifier les croyances des enseignants des cégeps à l'égard de l'évaluation des apprentissages en salle de classe;

2. Décrire les pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages et ce, en situation régulière de salle de classe.

Dans ce chapitre, seules les données quantitatives sont fournies. Les commentaires faits par les groupes réacteurs sont intégrés au volet d'analyse des résultats au chapitre suivant.

5.1.2 Composition des tableaux

Les tableaux de fréquences de l'échantillon total

Toutes les données recueillies sont traitées par le logiciel SPSS-Windows. Des tableaux de fréquences brutes et pondérées sont générés. Dans le texte qui suit, tous les tableaux sont construits à partir des proportions issues de fréquences pondérées. Toutes ces proportions décrivent l'échantillon total puisque nos répondants sont suffisamment nombreux pour représenter l'ensemble de la population.

Les tableaux d'analyses stratifiées

Pour chaque variable dépendante, l'analyse de la variance (ANOVA) est effectuée afin de discerner d'éventuelles différences significatives entre les strates. Un niveau de signification de 95% est adopté. Dans le présent chapitre, nous identifions, au fur et à mesure, les variables pour lesquelles des différences de moyennes entre les strates sont significatives à des seuils équivalents ou inférieurs à 0,05. Lorsqu'il y a différence significative entre les strates, nous décrivons et commentons les profils de réponses les plus éloignés de la moyenne afin de mettre en lumière les maxima et les minima et ce, en faisant référence à des tableaux que l'on retrouve en annexe 5. Dans ces tableaux, tous les choix de réponses ou tous les points sur les échelles de mesure ne sont pas nécessairement reproduits. Nous choisissons de présenter les colonnes qui contiennent les variations inter-strates les plus susceptibles d'illustrer les différences de profil entre les strates.

Lorsqu'il n'y a pas de différence significative entre les groupes, nous considérons qu'il y a tendance au consensus entre les strates et nous ne présentons pas d'analyses stratifiées. Les proportions de l'échantillon total sont alors considérées comme suffisantes.

Tous les tableaux de contingence qui sont présentés ont été validés selon les procédures admises (Everitt, 1988; Trudel et Antonius, 1991, p. 408-415). Le niveau de signification du χ^2 est d'au moins 95%. Les tableaux affichent la valeur "p" en plus du nombre "N" de répondants. On s'est assuré qu'aucune case vide (fréquences théoriques) ne soit présente et, si plus de 20% des cases ont des fréquences théoriques inférieures à 5, des colonnes ou des rangées sont regroupées.

Toutes les analyses stratifiées sont fondées sur un croisement entre la variable étudiée, d'une part, et la variable indépendante "strates", d'autre part. Cette variable indépendante est composée des dix (10) strates décrites au chapitre de la méthodologie ainsi qu'en annexe 4. En conséquence, la lecture des tableaux d'analyses stratifiées doit toujours se faire en gardant en mémoire que les appellations des strates correspondent à

des groupes de disciplines. Par exemple, la strate "Sciences humaines" inclut toujours la discipline Philosophie. De même, la strate "Lettres" inclut les Langues secondes.

Les appellations des strates sont:

CODES	Appellations complètes	Appellations abrégées
101	Sciences biologiques	Sc. biologiques
109	Éducation physique	Éduc. physique
199	Techniques biologiques	Tech. biologiques
201	Sciences physiques	Sc. physiques
299	Techniques physiques	Tech. physiques
301	Sciences humaines (incluant Philosophie)	Sc. humaines
399	Techniques humaines	Tech. humaines
400	Administration (pré-universitaire et technique)	Administration
500	Arts (pré-universitaire et technique)	Arts
601	Lettres	Lettres

Dans tous nos tableaux, seuls les pourcentages sont présentés et, en conséquence, toutes les comparaisons entre les groupes (strates) se font sur la base des proportions. Puisque les strates sont composées de nombres très inégaux de répondants, et que les strates échantillonnales ne sont d'ailleurs pas égales entre elles, au départ, toutes les strates sont ramenées à 100% et nous faisons des comparaisons inter-strates à partir de proportions obtenues par tableaux de contingence. Dans nos commentaires,

pour alléger la lecture, nous omettons l'expression "proportionnellement" mais il faut toujours la considérer comme implicite lorsque nous écrivons que les uns sont plus nombreux que les autres. Par exemple, "les professeurs de la strate x sont plus nombreux à afficher telle pratique que les professeurs de la strate y"; ou "c'est dans la strate x qu'on retrouve ce plus grand nombre de professeurs en accord". Dans ces deux exemples, nous comprenons que "les professeurs de la strate x sont proportionnellement plus nombreux que...".

5.1.3 Structure du chapitre

Toute la présentation des données est organisée en fonction de la typologie fondée sur les six champs de compétence que nous avons adoptés et décrits sommairement au chapitre 4 et repris en détail en annexe 7. Pour chaque champ de compétence, nous donnons, en alternance, les résultats des questions portant sur les croyances et, ensuite, les résultats des questions portant sur les pratiques. À la fin de cette séquence en douze parties, nous fournissons les résultats des questions portant sur la formation et les besoins de perfectionnement ainsi que les données de type démographique portant sur les répondants à notre questionnaire. Cependant, pour des raisons d'économie et de lisibilité, les très nombreux tableaux d'analyses stratifiées se retrouvent en annexe 5, dans le même ordre.

Les commentaires personnels recueillis à la question 67, lorsque pertinents, sont intégrés à l'analyse faite au chapitre 6.

Le tableau qui suit illustre la séquence de présentation des données et indique les numéros des questions appartenant à chaque cellule.

Tableau 3: Appartenance des énoncés du questionnaire à chacun des six champs de compétence

Les six champs de compétence (Catégories de Stiggins)	Les questions portant sur les Croyances	Les questions portant sur les Pratiques
1) L'utilisation de l'évaluation en classe;	Q6a à Q6k	Q4 et Q5
2) Les cibles de l'évaluation;	Q6l à Q6r	Q1 à Q3, Q43
3) Les qualités de l'évaluation;	Q33a à Q33j	Q7 à Q11, Q13 à Q20, Q22, Q24 à Q27, Q29, Q31, Q32
4) Les instruments d'évaluation;	Q33k à Q33s	Q12 et Q30; Q21 et Q34 à Q42
5) Les dimensions interpersonnelles de l'évaluation;	Q44a à Q44e	Q23, Q45 à Q48
6) La rétroaction et la notation.	Q49a à Q49i	Q28, Q50 à Q54
La formation et les besoins de perfectionnement	Q55 à Q58	
Les données de type démographique	Q59 à Q67	

5.2 Présentation des données

5.2.1 Première catégorie de Stiggins: L'utilisation de l'évaluation en classe

Selon les observations de Stiggins, en annexe 7, les enseignants se servent de l'évaluation des apprentissages pour répondre à trois catégories de besoins, à savoir: a) appuyer des décisions, b) guider l'enseignement et l'apprentissage, c) gérer la classe. Pour utiliser avec compétence l'évaluation des apprentissages dans le cadre de ces trois types de besoins, les professeurs doivent comprendre les fonctions de l'évaluation, ainsi que ses conséquences pédagogiques et éducatives, tant sur l'enseignement que sur l'apprentissage. Notre présentation suit l'ordre de ces trois types de besoins.

Les croyances relatives à l'utilisation de l'évaluation en classe

Appuyer des décisions

La grande majorité des professeurs croient que: 1) l'évaluation sert à identifier les forces et les faiblesses des étudiants en fonction des apprentissages à faire (Q6c), 2) suite aux résultats d'une évaluation des apprentissages, les enseignants devraient être disposés à réajuster leur contenu de cours (Q6d) et 3) l'évaluation ne sert pas principalement aux exigences administratives du collège (Q6i).

Par ailleurs, très peu de professeurs (14%) notent qu'ils n'évalueraient pas s'ils le pouvaient (Q6f). Le tableau C1-a, ci-après, illustre les pourcentages obtenus à chacun de ces énoncés. Deux points ressortent ici

- il y a quand même 14% des répondants (7% disent même être très en accord sur ce point) qui choisiraient de ne pas faire d'évaluation s'ils n'en étaient pas obligés.

- il y a près de 20% (un professeur sur 5) qui ne pensent pas que les professeurs devraient réajuster leur contenu si les résultats d'une évaluation le suggéraient. Peut-être ces professeurs ont-ils tendance à ne pas remettre en question le contenu de leurs cours et imputer l'éventuelle faiblesse de certains résultats d'évaluation à d'autres facteurs, soit au manque de préparation des étudiants à l'évaluation, soit à des défectuosités du dispositif d'évaluation lui-même. Sur cette question, il y a différence significative entre les strates. Le tableau C1-b, en annexe 5, montre que ce sont les professeurs de Techniques humaines qui sont les plus disposés à réajuster le contenu de leurs cours suite aux résultats d'une évaluation alors que ce sont les professeurs de sciences physiques qui le sont le moins.

Tableau C1-a: Distribution des réponses aux énoncés de croyances portant sur l'utilisation de l'évaluation en classe (énoncés abrégés)

		TA	A	D	TD	NSP
Guider des décisions						
Q6c)	L'évaluation sert à identifier les forces et les faiblesses de l'étudiant	43%	52%	4%	0%	1%
Q6d)	Suite à l'évaluation, les enseignants devraient être disposés à réajuster leur contenu de cours	32%	44%	15%	4%	4%
Q6f)	Si je le pouvais, je n'évaluerais pas	7%	7%	38%	45%	3%
Q6i)	L'évaluation sert principalement à satisfaire aux exigences administratives	2%	7%	53%	35%	3%
Aider à apprendre						
Q6e)	L'évaluation sert à vérifier ce que les étudiants ont appris	34%	61%	3%	1%	0%
Q6h)	L'évaluation peut aider les étudiants à apprendre	35%	60%	3%	0%	2%
Q6j)	N'est pas nécessaire de faire de l'évaluation formative	1%	3%	40%	51%	5%
Q6k)	L'évaluation formative, ce sont de petites évaluations qui servent à préparer à l'évaluation sommative	7%	43%	31%	13%	6%
Gérer la classe						
Q6a)	La meilleure façon de faire travailler les étudiants est de donner des points pour leur travail	17%	61%	17%	3%	2%
Q6b)	L'évaluation encourage les étudiants à fournir plus d'effort à l'étude	33%	61%	5%	0%	1%
Q6g)	Les étudiants travaillent plus régulièrement lorsque l'évaluation est fréquente	29%	58%	8%	1%	4%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération des strates. N min. = 616; N max. = 628

Aider à apprendre

La grande majorité des professeurs croient que l'évaluation sert à vérifier ce que les étudiants ont appris (6e) et qu'elle peut aider les étudiants à apprendre (6h).

Cependant, si la grande majorité est en désaccord à l'idée que l'évaluation formative n'est pas nécessaire au collégial (6j), on remarque qu'il y a différence significative d'opinion entre les strates sur ce point. Les professeurs de Techniques humaines sont les plus en désaccord alors que les professeurs de Sciences physiques sont les moins en désaccord (voir tableau C1-b, en annexe 5).

Quant à la conception que l'on se fait de l'évaluation formative, les opinions sont partagées assez fortement sur l'idée que l'évaluation formative constitue de "petites évaluations" qui servent à préparer à l'évaluation sommative (6k; tableau C1-a), mais il faut noter ici que 6% des répondants "ne savent pas" comment se positionner devant cette affirmation. Peut-être étaient-ils indécis devant cet énoncé parce qu'il n'était pas clair pour eux. Peut-être ne pouvaient-ils pas se prononcer sur l'évaluation formative comme concept? Ces deux hypothèses nous apparaissent vraisemblables et ce, pour deux raisons. D'une part, à la question précédente, 6j, traitant aussi d'évaluation formative, presque autant de répondants n'ont pas pu se prononcer (5% de "ne sais pas"). D'autre part, cette question ,6k, n'a pas fait l'objet d'une validation par les juges experts sur les critères de clarté et de pertinence. En effet, elle a été rédigée après le processus de validation, lors de la première mise en page, car nous avons alors estimé que le questionnaire ne comportait pas assez de questions sur l'évaluation formative.

Gérer la classe

Les professeurs reconnaissent que l'évaluation contribue à encourager les étudiants à fournir de l'effort à l'étude et au travail scolaire (6b). C'est là où il y a le moins d'indécis (1%) et personne n'est "très en désaccord".

Les croyances sont presque aussi fortes à l'égard de l'influence qu'exerce la fréquence de l'évaluation sur la régularité du travail scolaire (6g). Mais le fait de donner des points pour le travail scolaire est-il la meilleure façon de faire travailler les étudiants? Si la majorité croit que oui (6a), un professeur sur cinq est en désaccord (20%).

La mesure de chaque énoncé de croyance utilise une échelle qui permet au répondant d'indiquer, éventuellement, son indécision. Pour cette série de croyances relatives à l'utilisation de l'évaluation en classe, la proportion de "ne sais pas" atteint le seuil de 5% seulement pour les deux énoncés portant sur le concept d'évaluation formative:

5% des répondants ne savent que dire à l'énoncé 6j ("Au collégial, il n'est pas nécessaire de faire de l'évaluation formative.") et la proportion est de 6% pour l'énoncé 6k ("L'évaluation formative, ce sont de petites évaluations qui servent à préparer à l'évaluation sommative."). Les proportions de "ne sais pas" s'élèvent à 9% en Sciences humaines et en Sciences physiques pour le premier énoncé ainsi qu'en Sciences humaines et en Arts pour le second énoncé.

Les pratiques relatives à l'utilisation de l'évaluation en classe

Pour 81% des répondants, les travaux et les examens soumis à leurs étudiants contribuent "toujours" au calcul de la note inscrite au bulletin (Q4) alors que pour 17% des répondants, ils y contribuent "souvent" et pour 1%, "parfois". Ainsi, pour 98% des professeurs du collégial, les travaux et les examens soumis aux étudiants donnent lieu "toujours" ou "souvent" à une note sommative.

Nous avons également voulu connaître la proportion du temps de classe consacrée à des activités d'évaluation. La question 5 a été formulée de façon à ce que les répondants se fondent sur un format de cours équivalent à 45 heures et excluent les tâches effectuées à l'extérieur du collège (par exemple: les travaux réalisés à la maison).

Le tableau P1 illustre les proportions moyennes obtenues pour chaque choix de réponses. Puisqu'il y a différence significative entre les strates sur cette question, le tableau illustre aussi les proportions observées dans chaque strate. Les proportions les plus fortes dans chaque choix de réponses sont mises en évidence. On y constate que les professeurs d'éducation physique sont, proportionnellement, les plus nombreux à consacrer moins de cinq heures à des activités d'évaluation et ceux qui en consacrent le plus (plus de 18 heures), sont en Arts.

Tableau P1: Nombre d'heures, sur la base d'un cours de 45 heures, consacré à des activités d'évaluation

	Q5) Combien d'heures, sur un format de cours de 45 heures, vos étudiants sont-ils occupés à des activités d'évaluation de toute nature (exclure les travaux faits à l'extérieur du collège, tels les travaux réalisés à la maison)?					
STRATES:	**a) moins de 3 heures**	**b) de 3 à 5 heures**	**c) de 6 à 8 heures**	**d) de 9 à 12 heures**	**e) de 13 à 18 heures**	**f) plus de 18 heures**
Sc. biologiques	----	32%	32%	**24%**	3%	9%
Éduc. physique	**33%**	**50%**	13%	----	----	3%
Tech. biologiques	8%	36%	28%	14%	4%	10%
Sc. physiques	1%	26%	**47%**	16%	3%	6%
Tech. physiques	3%	35%	33%	14%	5%	10%
Sc. humaines	7%	24%	35%	16%	**13%**	4%
Tech. humaines	12%	37%	23%	17%	4%	8%
Administration	6%	14%	38%	22%	8%	12%
Arts	10%	22%	22%	14%	8%	**24%**
Lettres	5%	20%	43%	20%	5%	8%
Échantillon total:	**7%**	**27%**	**34%**	**17%**	**6%**	**10%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Ces chiffres obtenus pour l'échantillon total indiquent que, sur le "temps de classe", donc en excluant les travaux faits à l'extérieur:

- 34% des professeurs occupent leurs étudiants à des activités d'évaluation pour, approximativement, 10% et moins du cours (colonnes a) et b));
- 33% des professeurs occupent leurs étudiants à des activités d'évaluation pour, approximativement, 20% et plus du temps de cours (colonnes d), e) et f));

- 10% des professeurs, surtout en Arts, occupent leurs étudiants à des activités d'évaluation pour 40% et plus du temps de classe (colonne f)). Cela signifie qu'un professeur sur dix fait de l'évaluation des apprentissages, au moins 4 heures sur 10.

La formulation de la question fait référence à des activités d'évaluation de toute nature. La réponse de chacun inclut implicitement les évaluations sommatives, formatives et, s'il y a lieu, diagnostiques. Évidemment, pour chaque réponse, nous ne pouvons pas distinguer la proportion du temps de classe qui est consacrée à l'évaluation sommative et à l'évaluation formative.

5.2.2 Deuxième catégorie de Stiggins: Les cibles de l'évaluation

La seconde catégorie de Stiggins traite des objets sur lesquels portent l'évaluation. Les cibles généralement évaluées par les enseignants sont: la connaissance de la matière, le savoir-faire, les habiletés intellectuelles supérieures et les attitudes. Selon Stiggins, l'enseignant doit savoir comment évaluer chacune de ces cibles.

Les croyances relatives aux cibles de l'évaluation

La matière enseignée...

La majorité des répondants (58% et 57% respectivement) croient qu'il n'est pas correct de poser des questions d'examen au-delà de la matière enseignée (Q6m) et qu'il n'est pas utile que l'évaluation porte sur toute la matière enseignée (Q6n). De même, 62% des répondants sont en désaccord et très en désaccord avec l'idée que les examens devraient porter sur tout ce qui est dit en classe (Q6l) (voir tableau C2-a). Sur ce dernier point, il y a différence significative entre les strates . On note au tableau C2-b, en annexe, que les professeurs de Techniques biologiques sont les plus nombreux à être en désaccord (79%) alors que les professeurs de Sciences physiques et d'Arts (pré-universitaire et technique) sont les moins nombreux (52% et 54% respectivement).

Les habiletés à mesurer...

Quant aux habiletés qu'on peut ou devrait mesurer du niveau collégial, on observe que la grande majorité des répondants (68%) croient qu'il est plus important d'évaluer la compréhension de la matière que la connaissance des faits (Q6q). De même, ils ne croient pas qu'au collégial, il soit impossible d'évaluer autre chose que la connaissance de la matière (Q6o) et qu'il soit "pratiquement impossible" d'évaluer les attitudes (Q6p). Sur ces deux derniers points, il y a différence significative entre les

strates et le tableau C2-b illustre les variations. Notons ici que cette question de l'évaluation des attitudes a mené 13% des répondants à cocher "ne sais pas".

Cependant, les opinions sont très partagées quant à savoir si, au collégial, l'on devrait surtout évaluer les habiletés intellectuelles supérieures (analyse, synthèse, résolution de problèmes complexes) (Q6r). En effet, 50% des répondants sont en accord ou très en accord avec cet énoncé, 43% sont en désaccord ou très en désaccord et 7% ne "savent pas". Il y a, sur ce point, différence significative entre les strates (tableau C2-b). C'est en Sciences biologiques, en Sciences humaines et en Lettres que l'accord est le plus fort alors qu'il est le plus faible en Techniques physiques.

Tableau C2-a: Pourcentages des réponses aux énoncés (abrégés) de croyances portant sur les cibles de l'évaluation

		TA	A	D	TD	NSP
La matière enseignée						
Q6m)	**n'est pas** correct de poser questions d'examen au-delà de la matière enseignée	19%	39%	**33%**	5%	5%
Q6n)	**n'est pas** utile que l'évaluation porte sur toute matière enseignée	7%	50%	**28%**	**10%**	4%
Q6l)	examens devraient porter sur tout ce qui est dit en classe	8%	**27%**	50%	12%	3%
Les habiletés						
Q6q)	est plus important d'évaluer la compréhension que la connaissance	19%	**49%**	**20%**	4%	7%
Q6o)	est impossible d'évaluer autre chose que la connaissance	**2%**	**8%**	55%	32%	4%
Q6p)	pratiquement impossible d'évaluer les attitudes	5%	**19%**	46%	18%	**13%**
Q6r)	devrait surtout évaluer les habiletés intellectuelles supérieures	9%	**41%**	**40%**	4%	7%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération des strates. N min = 618; N max = 626

Les pratiques relatives aux cibles de l'évaluation

La première question posée aux enseignants nous apprend que, lors de la session d'hiver 1992, au moment où nous avons administré notre questionnaire, la moitié des répondants (51%) avaient l'intention de mesurer "tous les éléments du contenu" de leur cours. L'autre moitié des répondants se répartissent à peu près en parts égales, entre l'intention d'évaluer un échantillon du contenu (22%) et l'intention d'évaluer les

éléments essentiels seulement (27%). Le tableau P2-b, en annexe 5, montre que ce sont surtout les professeurs de la strate "Administration (pré-universitaire et technique)" qui ont l'intention de mesurer tout le contenu alors que les professeurs de la strate "Sciences biologiques" comptent ne mesurer qu'un échantillon et que les professeurs de la strate "Techniques physiques" ne mesureront que l'essentiel.

Au niveau des éléments mesurés (Q2), l'étude du mode nous donne quelques tendances révélatrices. La connaissance et la compréhension sont mesurées, le plus fréquemment, "toujours". Les habiletés d'application et d'analyse et synthèse sont mesurées, le plus fréquemment, "souvent". La réponse la plus fréquente, relativement à la mesure des habiletés de résolution de problèmes complexes, de la pensée critique et des attitudes, est "parfois". Les attitudes sont les moins mesurées. C'est là qu'on retrouve la plus grande proportion de "jamais" (17%) et de "rarement" (22%).

Pour certaines de ces habiletés mesurées, il y a des différences significatives entre les strates et le tableau P2-b, en annexe 5, illustre les moyennes obtenues de chaque strate, mettant en évidence les maxima et les minima parmi les strates.

On constate d'abord que parmi les éléments mesurés, seule "la compréhension de la matière" (Q2b) fait consensus entre les strates. C'est ici que la réponse "toujours" est la plus élevée: 61% des répondants disent mesurer toujours la compréhension. Quant aux autres apprentissages, le tableau P2-b illustre que:

- la connaissance est mesurée le plus souvent par les professeurs de Techniques humaines et le moins souvent en Arts (pré-universitaire et technique);
- l'application de procédures est mesurée le plus souvent en Éducation physique et en Arts (pré-universitaire et technique) et le moins souvent en Sciences biologiques;

- la résolution de problèmes complexes est mesurée le plus souvent en Arts (pré-universitaire et technique) et le moins souvent en Éducation physique et en Lettres;
- l'analyse et la synthèse est mesurée le plus souvent en Arts (pré-universitaire et technique) et en Lettres et le moins souvent en Éducation physique et en Administration (pré-universitaire et technique);
- la pensée critique est mesurée le plus souvent en Arts (pré-universitaire et technique) et le moins souvent en Sciences biologiques et en Éducation physique;
- les attitudes (savoir-être) sont mesurées le plus souvent en Éducation physique et en Arts (pré-universitaire et technique) et le moins souvent en Sciences physiques.

La question 3 présente une série de conditions et de critères auxquels les professeurs réfèrent pour attribuer des points. Ces conditions et critères sont, pour la plupart, extérieurs aux apprentissages attendus au collégial. Les réponses montrent qu'à divers degrés, ces conditions et critères sont tous incorporés aux notes cumulées et s'ajoutent à la note finale (tableau P2-a, ci-après). Mais pour chacune, il n'y a pas consensus entre les strates. Des différences significatives sont présentes entre les groupes, pour chacune de ces variables ($p < .05$). Le tableau P2-b, en annexe 5, présente ces sept variables dans l'ordre décroissant des réponses obtenues de l'ensemble des répondants et illustre les différences entre les strates, mettant en évidence les maxima et les minima.

Tableau P2-a: Proportions obtenues à la question 3: "Dans ce cours, en plus d'attribuer des points pour les objectifs liés à la matière enseignée, j'en accorde pour:"

Rang	Énoncés	Échantillon total
1	g- le respect de mes exigences de présentation des travaux	71%
2	c- l'exécution de devoirs ou d'exercices à la maison ou en classe	69%
3	f- la créativité, l'originalité	53%
4	a- la motivation et l'effort dans l'apprentissage	48%
5	e- la participation active de l'étudiant en classe	40%
6	b- la présence assidue aux cours	37%
7	d- les progrès réalisés au cours de la session (évolution)	36%

Note: Calculs sur fréquences pondérées

Quant au choix des cibles de l'évaluation, on observe que la majorité des professeurs (55%) décident seuls de "ce sur quoi doit porter (leur) évaluation" (Q43). Dans le tableau qui suit, nous présentons les données en parallèle avec les réponses à la question 42, relative à l'instrumentation.

Tableau P2-a: (suite) Qui fait les choix en matière de...

	Q42) Instruments d'évaluation	Q43) Cibles de l'évaluation
Le professeur seul	44%	55%
Le professeur et les collègues donnant le même cours	39%	34%
Le professeur et son département	13%	7%
Le professeur et ses étudiants	3%	3%

Note: Calculs sur fréquences pondérées

5.2.3 Troisième catégorie de Stiggins: Les qualités de l'évaluation

Les caractéristiques d'une "bonne" évaluation varient selon le contexte de l'évaluation. Cependant, quelques critères de qualité sont communs à toutes les situations d'évaluation: la correspondance entre l'objet à évaluer et le dispositif de mesure choisi; le contrôle des sources d'erreur de mesure; la concordance entre l'apprentissage visé et les résultats d'évaluation obtenus; l'obtention d'une information dont l'interprétation est claire pour l'enseignant et les étudiants.

Les qualités de l'évaluation ne s'expriment donc pas seulement en termes de fidélité et de validité des instruments de mesure. Les enseignants doivent savoir résoudre des problèmes relatifs à la conception, à l'administration et à la notation des dispositifs de mesure ainsi qu'à l'interprétation des notes et à l'utilisation qui en sera faite.

Les croyances relatives aux qualités de l'évaluation

La qualité des items et des consignes

Trois énoncés mesurent les croyances des répondants relativement à certaines pratiques susceptibles de favoriser la qualité des items et des consignes. La grande majorité croient que (tableau C3-a):

- après la passation d'un examen, il est utile de faire l'analyse des questions pour en évaluer la qualité (Q33a);
- les questions d'examen ont davantage de valeur lorsqu'elles sont lues par un second spécialiste de la matière (Q33b);
- on s'assure davantage de la clarté des consignes des travaux écrits en les faisant examiner par un second spécialiste de matière (Q33c).

Mais, sur les deux derniers énoncés, les pourcentages de désaccord sont supérieurs à 10%. De plus, il y a différence significative entre les strates sur chacun de ces trois énoncés. Le tableau C3-b, en annexe 5, illustre les proportions de réponses par strate et montre les maxima et minima par énoncé. On y voit que, dans l'ensemble des trois énoncés, les proportions de désaccord ou de "ne sais pas" sont les plus élevées dans les strates Arts (pré-universitaire et technique) ou Sciences humaines (incluant Philosophie). À l'opposé, ce sont les professeurs de Techniques biologiques qui sont le plus en accord avec les trois énoncés.

Les critères d'évaluation

Deux énoncés, 33d et 33i, réfèrent aux critères d'évaluation. Dans ces deux cas, plus de 90% des répondants croient que tous les étudiants d'un même cours devraient être évalués à partir des mêmes critères (33d) et qu'il est nécessaire d'établir les critères d'évaluation avant la correction (33i). Cependant, il y a différence significative entre les strates sur ce dernier point. Près d'un professeur sur dix, en Sciences biologiques, ne croit pas qu'il soit nécessaire d'établir les critères d'évaluation avant la correction, alors que la strate Techniques humaines affiche la proportion la plus forte de "très en accord" (tableau C3-b).

Représentativité et validité

Les questions 33e, f, g et j mesurent les croyances des professeurs sur divers aspects de la validité des évaluations.

La grande majorité des répondants de toutes les strates pensent qu'il faut faire plusieurs évaluations pour pouvoir se fier davantage à l'ensemble des résultats obtenus (33f), que "mes moyens d'évaluation me disent ce que je veux savoir" (33g), et que les notes devraient refléter le nombre ou le pourcentage des objectifs maîtrisés par les étudiants (33j). Dans ce dernier cas, cependant, il y a différence significative entre les strates: les

professeurs de Sciences humaines (incluant Philosophie) sont les plus nombreux à être en désaccord ou à répondre "ne sais pas" à l'énoncé disant que les notes devraient refléter le nombre ou le pourcentage des objectifs maîtrisés par les étudiants (tableau C3-b).

Quant à la croyance à l'effet qu'il est pratiquement impossible de faire en sorte que les résultats des évaluations représentent bien les apprentissages des étudiants (33e), la majorité la rejette, mais il faut noter qu'il y a quand même 29% des répondants, sans différence entre les strates, qui sont en accord avec cet énoncé.

Un dernier énoncé (33h) mesure la perception de rigueur dans les pratiques d'évaluation en éducation. Cette question sollicite manifestement, chez nos répondants, un jugement de valeur. Pourtant, 98% des répondants ont choisi de répondre à cette question. De ceux-là, 38% ont indiqué "ne sais pas" alors que 40% ont signifié que beaucoup d'enseignants manquent de rigueur à l'égard de l'évaluation (33h) et 22% se sont dits en désaccord avec l'énoncé.

Tableau C3-a: Réponses aux énoncés de croyances portant sur les qualités de l'évaluation

		TA	A	D	TD	NSP
Validation des items						
Q33a)	Il est utile de faire l'analyse des questions d'examen	35%	**56%**	4%	1%	3%
Q33b)	Les questions d'examen ont + de valeur si lues par un second spécialiste	14%	**58%**	15%	2%	**11%**
Q33c)	Les consignes des travaux écrits sont plus claires si examinées par un second spécialiste	17%	**64%**	11%	1%	**7%**
Critères d'évaluation						
Q33d)	Tous devraient être évalués à partir des mêmes critères	47%	45%	5%	1%	2%
Q33i)	Nécessaire d'établir les critères avant la correction	47%	48%	4%	----	1%
Représentativité						
Q33e)	Il est impossible que les résultats des évaluations représentent bien les apprentissages des étudiants	4%	25%	**50%**	16%	5%
Q33f)	Faut plusieurs évaluations pour se fier aux résultats	32%	**62%**	5%	----	1%
Q33g)	Mes moyens d'évaluation sont fiables	13%	**79%**	4%	----	4%
Q33j)	Les notes devraient refléter les objectifs maîtrisés	22%	**62%**	8%	1%	**7%**
Rigueur						
Q33h)	Beaucoup d'enseignants manquent de rigueur en évaluation	**11%**	**30%**	20%	2%	**38%**

Note: Ces % proviennent des tableaux de fréquences calculées après pondération des strates. N min = 613; N max = 618

Les pratiques relatives aux qualités de l'évaluation

Planification de l'évaluation

Lorsqu'ils rédigent un instrument de mesure, on constate que la majorité des répondants commencent par faire un plan détaillé de ce qu'ils veulent évaluer (Q8). Par ailleurs, 8% disent ne le faire que rarement ou jamais (voir le tableau P3-a).

Très peu de professeurs disent avoir investi plus de 12 heures à planifier et à rédiger des dispositifs d'évaluation sommative (Q10) dans la période de deux semaines coïncidant avec notre questionnaire:

- moins de 3 heures: 40%
- de 3 à 6 heures: 43%
- de 7 à 9 heures: 10%
- de 10 à 12 heures: 4%
- plus de 12 heures: 3%

La lecture de ces réponses à la question 10 exige évidemment de la prudence. Il n'y a pas de différence significative entre les strates mais, d'une strate à l'autre, les durées de cours varient. Les cours d'éducation physique sont généralement de deux heures par semaine alors que les cours de soins infirmiers, par exemple, peuvent demander, à certaines périodes d'une session, plus de 15 heures par semaine pour un même groupe-cours.

Dans une perspective de planification de l'évaluation, beaucoup d'enseignants prétendent réutiliser des questions d'examen d'une session à l'autre, soit en empruntant des questions à des collègues, soit en se donnant des banques de questions d'examen. Cependant, 22% des enseignants ne réutilisent que très rarement, sinon jamais des questions déjà existantes, et près de 40% n'y ont recours que parfois (tableau P3-a). Il y a différence entre les strates sur cette question. On constate que cette

pratique est plus fréquente en Sciences physiques et moins fréquente en Lettres.

Contrôle de la qualité, a priori

Seuls 5% des répondants utilisant l'examen disent faire toujours examiner leurs questions par un ou des collègues avant de l'administrer (Q13) alors que 27% le font rarement, 28% le font parfois et 24% ne le font jamais.

Des répondants utilisant le travail écrit comme moyen d'évaluation, seuls 7% font toujours examiner leurs consignes de travaux écrits (Q31) par un ou des collègues avant de les présenter aux étudiants alors que 29% le font parfois, 24% rarement et 20% jamais.

De façon générale, il semble qu'au collégial, la préparation des examens est un travail plutôt solitaire.

Sur ces deux questions, il y a différence significative entre les strates. On constate que les professeurs de Techniques biologiques réfèrent à leurs collègues plus souvent alors que les professeurs de Techniques physiques, de Sciences humaines (incluant Philosophie) et de Lettres le font moins fréquemment.

Contrôle de la validité

Afin de décrire en quoi les professeurs s'assurent de la validité de contenu de leurs examens, deux questions mesuraient les pratiques qu'ils mettent en oeuvre pour vérifier s'il y a correspondance entre leurs examens d'une part et leur contenu ou leurs objectifs d'autre part. Ce sont les questions 16 et 17.

Pour ces deux questions, il n'y a pas de différence significative entre les strates. En effet, le profil de réponses est semblable pour l'ensemble des répondants. Par contre, lorsqu'on compare les réponses aux deux questions entre elles, on constate que:

- dans l'ensemble, les professeurs se servent d'abord et principalement de leurs objectifs pour préparer la plupart de leurs questions d'examen. Pour en donner une mesure, notons que, de tous ceux qui travaillent toujours à partir de leurs objectifs, 65% se réfèrent toujours aussi à leur contenu;

- les strates Sciences humaines (incluant Philosophie) ainsi que Administration (pré-universitaire et technique) semblent afficher des profils semblables aux deux questions. Les professeurs de ces strates disent préparer la plupart de leurs questions d'examen autant en relisant le contenu de leurs cours qu'à partir des objectifs de leurs cours.

Fréquence de l'évaluation sommative

Quant à la fréquence de l'évaluation (Q11), la réponse la plus fréquente, le mode, nous apprend que les professeurs font généralement de l'évaluation sommative aux trois semaines (46%). Les professeurs de Techniques biologiques sont les plus nombreux à participer à ce mode, ce qui signifie que ce sont eux qui, en moyenne, évaluent le plus souvent "aux trois semaines". Le tableau P3-b, en annexe 5, montre que parmi ceux qui font de l'évaluation sommative hebdomadaire, on retrouve surtout les professeurs de Lettres et d'Arts. À l'inverse, c'est en Éducation physique qu'on trouve la plus grande proportion évaluant "aux 5 semaines ou plus rarement".

Tableau P3-a: Les pratiques relatives aux qualités de l'évaluation. Réponses de l'échantillon total

		T	S	P	R	J
Planification de l'évaluation						
Q8)	Je fais un plan détaillé	38%	39%	16%	5%	3%
Q15)	J'utilise des questions déjà existantes	4%	35%	39%	17%	5%
Contrôle de la qualité, a priori						
Q13)	Je fais examiner mes questions par un ou des collègues avant d'administrer un examen	5%	17%	28%	27%	24%
Q31)	Je fais examiner mes consignes de travaux écrits par un ou des collègues avant de les présenter aux étudiants	7%	21%	29%	24%	20%
Contrôle de la validité						
Q16)	Je relis le contenu de mon cours	36%	47%	11%	4%	1%
Q17)	Je me fonde sur les objectifs du cours	48%	40%	8%	4%	----
Administration de l'évaluation						
Q20a)	J'écris les questions au tableau ou sur acétates	----	2%	3%	6%	88%
Q20b)	Je les imprime et distribue	91%	7%	1%	----	----
Q20c)	Je fais la lecture des questions à haute voix	3%	2%	2%	5%	88%

Administration de l'évaluation (questions 20-24-25-26)

Dans la très grande majorité des cas (92%), les professeurs administrent toujours leurs questions d'examen en les imprimant et en distribuant un exemplaire à chaque étudiant (Q20). De même, ils ne vont presque jamais écrire les questions au tableau ou les projeter sur acétates, ou en faire la lecture à haute voix afin que les étudiants les prennent en note. Les exceptions sont rares. Seulement 5% des 460 répondants à ces questions écrivent "souvent" ou "parfois" leurs questions au tableau (ou les projètent sur acétates). On remarque qu'il y a 14% de "souvent" dans la strate "Lettres" (voir tableau P3-b). Deux répondants disent même faire cela "toujours".

D'autre part, 6% des 460 répondants disent qu'ils font "toujours", "souvent" ou "parfois" la lecture à haute voix des questions d'examen. Le tableau P3-b montre qu'il y a 12% de "toujours" dans la strate Arts et 13% de "parfois" dans la strate Lettres.

Les examens à développement sont des dispositifs d'évaluation très largement utilisés au collégial. Grâce à la question 21, traitée plus loin (catégorie 4: l'instrumentation de l'évaluation), nous savons que les questions à développement court sont "toujours" utilisées par un professeur sur quatre, ce qui est supérieur à l'usage des questions à choix multiples.

Chez les professeurs qui emploient des questions à développement (Q24), trois professeurs sur cinq (60%) proposent aux élèves, entre rarement et toujours, des questions "au choix" (par exemple: "répondre à 3 questions parmi un choix de 8"). Cette pratique se manifeste dans toutes les strates, à des degrés divers, mais le plus fortement en Sciences humaines (incluant Philosophie).

Par ailleurs, il arrive à moins de 25% des répondants de permettre parfois ou rarement aux étudiants de répondre à leurs examens "à la maison" (Q25) et de remettre leurs réponses lors de la prochaine rencontre. Une analyse croisée des réponses aux deux questions 24 et 25 montre que, chez

ceux qui proposent toujours ou souvent des questions "au choix", 40% et 18%, respectivement, permettent parfois aux étudiants de répondre "à la maison".

Enfin, si le tiers des professeurs, et tout particulièrement en Sciences biologiques et en Techniques biologiques, ne permettent jamais aux étudiants d'avoir en main toute documentation utile lors d'examens composés de questions à développement (Q26), les professeurs de la strate Lettres semblent les plus nombreux à accorder cette permission. Le tableau P3-b, en annexe 5, illustre les minima et maxima pour ces trois dernières questions.

Plusieurs auteurs croient qu'au niveau collégial, les étudiants devraient avoir accès à toute documentation utile. Ils pensent qu'à ce niveau, on devrait davantage demander aux étudiants de démontrer s'ils sont capables de résoudre des problèmes nouveaux, de construire des synthèses et de faire des analyses plutôt que de rendre compte des connaissances mémorisées. Pour eux, les dispositifs de mesure doivent être construits de façon à mesurer ces habiletés et non pas la capacité de se rappeler des données.

Pour notre part, nous pensons que la distribution des pratiques à la question 26 reflète probablement la diversité des réalités disciplinaires. Certains cours sont fondés sur de grandes quantités de connaissances déclaratives: concepts, vocabulaire, procédures, etc. Et c'est possiblement le cas dans beaucoup de disciplines pour lesquelles les étudiants en sont à leur premiers contacts ou pour lesquelles ils sont insuffisamment préparés.

Tableau P3-a: (suite) Les pratiques relatives aux qualités de l'évaluation. Réponses de l'échantillon total

		T	S	P	R	J
Dans les questions à développement...						
Q24)	Propose des questions "au choix"	4%	11%	18%	20%	32%
Q25)	Permet de répondre à la maison	----	1%	9%	11%	62%
Q26)	Permet toute documentation	9%	13%	21%	13%	29%
Préparation des étudiants						
Q18)	Préviens à l'avance du contenu de l'examen	75%	17%	4%	2%	1%
Q19a)	Fais faire des exercices équivalents	18%	37%	22%	13%	10%
Q19b)	Fais rédiger des questions d'examen	----	1%	7%	16%	76%
Q19c)	Fais démonstrations et révisions	20%	38%	25%	10%	7%
Q32a)	J'explique oralement la tâche	71%	20%	4%	2%	2%
Q32b)	Remets une feuille explicative	58%	22%	13%	4%	3%
Q32c)	Étapes de rétroaction corrective, en route	18%	26%	26%	17%	12%
Correction et notation						
Q9)	Donne les critères de correction à l'avance	41%	34%	14%	7%	3%
Q27)	Pénalise le "guessing"	5%	4%	8%	8%	44%
Améliorations a posteriori						
Q14)	Fais l'analyse des items	26%	40%	25%	6%	2%
Q22)	Demande l'opinion des étudiants sur les items	6%	22%	26%	27%	18%
Q29)	Conserve mes questions pour les réutiliser	40%	34%	20%	5%	1%

Note: Calculs fondés sur les fréquences pondérées

Préparation des étudiants à répondre aux examens

Les répondants du réseau disent contribuer à aider leurs étudiants à se préparer à répondre aux questions d'examen. En effet, les trois quarts des professeurs préviennent toujours les étudiants, à l'avance, de ce sur quoi portera l'examen (Q18). Un autre 18% le font souvent. Seulement 3% des professeurs ne le font que rarement sinon jamais et ce, sans distinction entre les strates.

De plus, ils mettent en oeuvre diverses stratégies pour aider les étudiants à se préparer à répondre aux questions d'examen (Q19). Un peu plus de la moitié des répondants font faire ou proposent des exercices avec des questions équivalentes (Q19a) ou encore font des démonstrations ou des révisions de la matière (Q19c) et ce, souvent ou toujours. Mais près d'un professeur sur cinq ne proposent des exercices équivalents que rarement, sinon jamais, 15% font des démonstrations ou des révisions de la matière rarement ou jamais (voir tableau P3-a).

Enfin, plus des trois quarts des répondants ne font jamais rédiger une partie des questions d'examen par les étudiants eux-mêmes (Q19b) et 16% utilisent rarement cette stratégie pour aider leurs étudiants à se préparer à répondre aux examens.

Préparation des étudiants à exécuter des travaux écrits

Quant à savoir comment les professeurs aident leurs étudiants à exécuter les travaux écrits qui contribuent au calcul de la note finale (Q32), on remarque que:

- l'immense majorité explique la tâche oralement (Q32a) souvent (20%), voire toujours (71%);
- une autre majorité, moins forte, remet une feuille explicative aux étudiants (Q32b): 58%, toujours et 22% souvent.

Enfin, une troisième stratégie pour aider les étudiants à exécuter les travaux écrits, est de prévoir des étapes de rétroaction corrective avant la date de remise du travail (Q32c). Cette pratique est beaucoup plus rare car 30% des professeurs n'en prévoient jamais, sinon rarement et 26% en prévoient parfois. Et ces chiffres valent pour l'ensemble des répondants car, sur ce point, il n'y a pas de différence significative entre les strates.

Correction et notation

Lors de la correction d'un examen ou d'un travail (Q7), les professeurs se fondent généralement sur des critères d'évaluation déterminés à l'avance (83%) ou tout juste avant le début de la correction (11%). Mais 7% des professeurs attendent le moment de la correction pour déterminer leurs critères: 2% le font "au cours de la correction" et 5% font d'abord une première lecture des copies des étudiants. Le tableau P3-b, en annexe 5, montre les différences entre les strates.

La proportion des professeurs qui déterminent les critères de correction à l'avance est la plus forte en Techniques biologiques et la moins forte en Sciences humaines (incluant Philosophie), alors que c'est dans ce dernier groupe qu'on retrouve, à l'inverse, la plus forte proportion de critères déterminés "après une première lecture des copies des étudiants".

Ces critères sont communiqués aux étudiants à l'avance (Q9) pour la plupart des professeurs, mais plus d'un professeur sur dix ne le fait que rarement (8%) sinon jamais (3%). Ici, les professeurs de Sciences physiques sont proportionnellement les plus nombreux à ne jamais communiquer les critères à l'avance, alors que les professeurs de la strate Techniques humaines sont les plus nombreux à toujours le faire à l'avance.

Pour compenser la possibilité que les étudiants choisissent leurs réponses au hasard lorsqu'ils ne connaissent pas la réponse à une question objective, il existe des méconnaissances permettant de corriger les effets du

hasard sur les résultats (Morissette, 1993). Chez les professeurs qui utilisent des questions à choix multiples, 63% ne font jamais ces calculs correctifs. Moins d'un professeur sur dix (7%) se prévaut toujours d'un système de pénalité pour compenser le choix au hasard. Cette pratique est la plus fréquente chez les professeurs de Sciences physiques, la moins fréquente en Sciences biologiques et en Techniques biologiques.

Amélioration a posteriori (questions 14-22-29)

Trois de nos questions cherchent à mesurer les pratiques des professeurs menant à améliorer, a posteriori, leurs dispositifs d'évaluation à partir des informations obtenues lors de l'administration de ces dispositifs.

Le tableau P3-a montre que près de 7 professeurs sur dix (Q14) font toujours ou souvent, après la correction d'un examen, l'analyse de leurs questions d'examen (clarté, compréhension, difficulté, plausibilité des choix de réponses s'il y a lieu, etc.). Il faut cependant noter que, sur l'ensemble des répondants, 8% ne font cette analyse que rarement, sinon jamais. Le tableau P3-b illustre les différences par strates et met en évidence les minima et maxima.

Il arrive peu fréquemment que les répondants demandent aux étudiants leur opinion sur la formulation des items du questionnaire lorsque les résultats d'un examen sont remis (Q22). En effet, 17% ne sollicitent jamais de rétroaction de la part de leurs étudiants et 29% ne le font que rarement et ce, surtout en Sciences biologiques, en Arts (pré-universitaire et technique) et en Lettres.

Cependant, la conservation des questions d'examen en vue de les réutiliser (Q29) s'avère une pratique plus répandue: près du trois-quart des professeurs le font toujours ou souvent. Seulement 7% des professeurs ne conservent jamais leurs questions ou ne le font que rarement. C'est le cas particulièrement en Arts (pré-universitaire et technique) et en Lettres.

5.2.4 Quatrième catégorie de Stiggins: Les instruments d'évaluation

Selon Stiggins, les professeurs utilisent au moins trois formes d'évaluation des apprentissages en salle de classe: les instruments de type "papier-crayon", l'observation et l'échange verbal. Les professeurs doivent savoir se servir de ces instruments tout en comprenant que les règles de validité sont différentes d'une forme à l'autre.

Les croyances relatives aux instruments

Examens et questions d'examens

Quatre des six questions de cette série mesurent les croyances des professeurs à l'égard des questions d'examens, souvent appelées items. Pour ces quatre questions, il y a différence significative entre les strates et les tableaux C4-b, en annexe 5, sont conçus pour illustrer les maxima et minima.

La majorité des répondants ne croient pas que seules les questions à développement peuvent mesurer les apprentissages de haut niveau cognitif: application, analyse, synthèse (Q33k). Mais à la question 33r (tableau C4-a, ci-après), on voit qu'une majorité ne croit pas non plus que les questions à choix multiples peuvent mesurer les apprentissages de haut niveau cognitif (application, analyse, synthèse). Pour ces deux questions, les majorités sont faibles car les proportions sont presque équivalentes entre les répondants en accord et les répondants en désaccord. Et les professeurs qui "ne savent pas" restent près du 10%. Le tableau C4-b illustre ces différences et montre que, si les professeurs de Lettres et de Sciences biologiques sont les plus nombreux à favoriser les questions à développement pour mesurer les apprentissages de haut niveau, ce sont les professeurs de Techniques biologiques qui, les plus nombreux, accordent ce pouvoir aux questions à choix multiples. À l'inverse, si les professeurs de Techniques biologiques et en Administration (pré-universitaire et technique) favorisent le moins les questions à développement,

ce sont les professeurs d'Arts (pré-universitaire et technique) qui font le moins confiance aux questions à choix multiples.

Si 40% des répondants, principalement en Techniques humaines et en Lettres, croient que les questions à choix multiples mesurent principalement les connaissances mémorisées (Q33o), 50% des professeurs, tout particulièrement dans les strates Techniques biologiques et Sciences physiques, ne croient pas que ce soit le cas. Le tableau C4-b illustre ces croyances et montre les minima et maxima.

Tableau C4-a: Pourcentages des réponses aux énoncés de croyances portant sur les instruments d'évaluation

		TA	A	D	TD	NSP
Q33k)	Les questions à développement peuvent seules mesurer les apprentissages de haut niveau	12%	32%	**41%**	7%	8%
Q33r)	Les questions à choix multiples peuvent mesurer des apprentissages de haut niveau	3%	38%	**32%**	**15%**	**11%**
Q33o)	Les questions à choix multiples mesurent principalement les connaissances	6%	35%	**42%**	8%	9%
Q33s)	Les questions à choix multiples font l'objet de réponses au hasard trop souvent	8%	**35%**	33%	2%	**21%**
Q33l)	Examens à livre ouvert	4%	14%	**55%**	12%	**15%**
Q33p)	Examens-synthèse obligatoires dans toutes disciplines	13%	**38%**	29%	5%	**15%**
Q33q)	Travaux écrits et projets sont les meilleures évaluations	6%	31%	**44%**	5%	13%
Les évaluations ... devraient surtout être des autoévaluations:						
Q33m)	Formative	4%	24%	**48%**	13%	**10%**
Q33n)	Sommative	0%	5%	**53%**	**35%**	6%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération des strates. N min = 607; N max = 619

Comportements des étudiants

À l'énoncé 33s, on note que l'attitude des professeurs à l'égard des questions à choix multiples se maintient. Les professeurs de Techniques biologiques sont ceux qui sont le moins en accord avec l'idée que les étudiants répondent trop souvent au hasard à des questions à choix multiples alors que les professeurs d'Arts (pré-universitaire et technique) et de Lettres demeurent ceux qui manifestent le plus fréquemment leur accord avec cet énoncé. Le tableau C4-a, ci-après, montre que 21% des répondants "ne savent pas". Cette forte proportion peut s'expliquer par le fait que l'énoncé 33s réfère à une problématique très technique pour laquelle plusieurs professeurs n'ont pas d'opinion.

La majorité des répondants (67%) ne croient pas que les examens devraient être, de préférence, à livre ouvert (Q33l). Une remarque s'impose ici. Un collaborateur nous a signalé, à l'époque où les questionnaires ont été soumis à l'échantillon d'enseignants, que l'expression "livre ouvert" pourrait comporter un risque de confusion chez le répondant. Selon ce collaborateur, le dictionnaire Robert associe l'expression "livre ouvert" à l'adverbe "couramment" (Petit Robert, au mot livre). Ainsi, un répondant pourrait lire l'énoncé 33l ainsi: "les examens devraient être habituels, ordinaires". Si cette compréhension est jugée vraisemblable, notre énoncé ne serait plus univoque et les statistiques obtenues deviennent ininterprétables.

Depuis quelques années, plusieurs débats ont eu lieu sur la pertinence ou la nécessité de prévoir des examens-synthèse à la fin des sessions. L'énoncé 33p nous apprend que 50% des répondants croient que des examens-synthèse de fin de session devraient être obligatoires dans presque toutes les disciplines alors que 34% disent être en désaccord et que 15% ne savent pas. Le taux de "ne sais pas" est particulièrement élevé sur cette question, reflétant peut-être les opinions souvent contradictoires qui s'expriment lors des débats évoqués ci-haut.

Il n'y a pas de différence significative entre les strates sur cet énoncé ($p > .05$) indiquant que les fréquences observées (tableau C4-a) ont tendance à refléter des consensus entre les strates.

Les productions écrites

Pour compléter ces énoncés portant sur diverses facettes de l'examen et de certains types d'items, l'énoncé 33q porte, quant à lui, sur une autre "famille" de dispositifs d'évaluation: les productions étudiantes.

Les professeurs croient peu que la rédaction de travaux écrits ou la réalisation de projets sont les meilleurs moyens d'évaluation (Q33q). Parmi ceux qui y croient le moins, on retrouve les professeurs de Techniques biologiques et de Sciences physiques. À l'inverse, ceux qui ont le plus tendance à être en accord avec l'énoncé sont les professeurs en Arts et en Lettres.

On constatera que ces croyances correspondent, avec une corrélation forte, aux pratiques des professeurs. Si les professeurs ont des croyances favorables envers un dispositif de mesure, on constatera qu'ils en font usage.

L'autoévaluation

Si l'évaluation peut être faite au moyen d'examens ou de travaux écrits, certains préconisent de confier aux étudiants eux-mêmes l'évaluation de leurs apprentissages. Le recours à l'autoévaluation peut être invoqué tant en évaluation formative que sommative et notre questionnaire propose deux énoncés distincts pour mesurer les croyances à l'égard de l'un et de l'autre.

Si la majorité des répondants ne croient pas que les évaluations devraient surtout être des autoévaluations, tant en fonction formative qu'en fonction sommative, il reste que plus d'un professeur sur quatre croit que les

évaluations formatives devraient surtout être des autoévaluations (Q33m).

À cet égard, il y a différence entre les strates et le tableau C4-b, en annexe 5, montre que la strate Sciences biologiques compte la proportion la plus élevée de répondants en accord avec cette idée. À l'opposé, les professeurs des strates Éducation physique, Arts (pré-universitaire et technique) et Sciences humaines (incluant Philosophie) sont proportionnellement les plus nombreux à être en désaccord sur ce point.

Lorsque l'évaluation a une fonction sommative, il y a beaucoup plus facilement consensus entre les strates: 88% sont en désaccord sinon très en désaccord à l'idée que les évaluations sommatives devraient surtout être des autoévaluations (Q33n) et ce, sans différence significative inter-strates (tableau C4-a).

Les pratiques relatives aux instruments d'évaluation

Deux questions-filtres, les questions 12 et 30, permettent aux répondants de répondre ou d'éviter de répondre à des groupes de questions relatives à deux catégories de dispositifs d'évaluation, à savoir, les examens et les travaux écrits. En même temps, ces questions-filtres nous permettent de dresser un tableau des instruments qu'utilisent les professeurs. On y constate, notamment, que l'examen et le travail écrit sont utilisés par 89% et 71%, respectivement, des répondants au questionnaire. C'est en Sciences physiques, en Techniques physiques et en Administration (pré-universitaire et technique) qu'on utilise l'examen le plus alors que c'est en Arts (pré-universitaire et technique) qu'on l'utilise le moins. Quant au travail écrit, on l'utilise le plus en Lettres et en Techniques humaines et le moins en Sciences physiques (tableau P4-b).

Items

Ceux qui utilisent l'examen pour effectuer l'évaluation de leurs étudiants peuvent faire appel à deux genres d'items. Il s'agit des questions objectives, dans lesquelles l'étudiant choisit la bonne ou la meilleure réponse, et les questions à développement, qui demandent à l'étudiant de construire sa réponse. Parmi les questions à développement, on distingue le développement long et le développement court (Morissette, 1993).

La question 21 nous apprend que ce sont les questions à développement court qui sont les plus utilisées, toujours ou souvent, principalement en Sciences biologiques et en Techniques physiques.

En second lieu, les questions objectives à choix multiples sont utilisées, toujours ou souvent, par 42% des répondants dont, principalement, les professeurs en Éducation physique et Techniques biologiques.

Enfin, les questions à développement long sont utilisées par 33% des professeurs, souvent ou toujours.

Le tableau P4-b montre quelques tendances. On y lit que les professeurs de Techniques biologiques sont les plus nombreux, proportionnellement, à ne jamais utiliser les questions à développement, longues ou courtes, et à toujours utiliser les questions objectives, à choix multiples. De même, c'est dans la strate Lettres qu'on retrouve la plus grande proportion de "jamais" pour les questions objectives et la plus forte proportion de "souvent" pour les questions à développement long.

Instrumentation

La question 34 propose au répondant une série d'approches pouvant servir à l'évaluation des apprentissages au collégial. La liste de ces approches nous a été suggérée par les interviews de pré-consultation ainsi que par la revue de la documentation. Dans l'énoncé de la question, nous avons choisi l'expression "approches" plutôt que "dispositifs

d'évaluation" car, si nous voulons mesurer l'utilisation des devoirs ou des exercices en classe à des fins d'évaluation des apprentissages en classe, nous ne voulons pas tromper le lecteur en associant tous les objets de notre liste à des dispositifs d'évaluation.

Il serait oiseux de présenter, strate par strate, les réponses obtenues à chacun des cinq points de l'échelle car les profils de réponses varient beaucoup, à l'image de la réalité disciplinaire de chaque strate. Le tableau P4-a, ci-après, décrit l'utilisation qui est faite des dix-sept approches proposées en définissant l'utilisation par la somme des "toujours" et des "souvent" à la question 34. En annexe 5, le tableau P4-b reprend chacune des dix-sept approches en identifiant les strates qui en font usage selon des proportions égales ou supérieures à la moyenne observée dans l'échantillon total. Seul "l'examen surprise" (Q34m) n'y est pas présenté, car aucune strate n'en fait usage selon une proportion supérieure à la moyenne, déjà très faible (3%). Dans la perspective du troisième objectif de cette recherche qui consiste à dégager des pistes de perfectionnement, cette série de tableaux sert à identifier les strates qui font usage des approches d'évaluation et qui constituent des clientèles-cibles d'offres de perfectionnement.

Les questions 35 à 41 servent à montrer comment les professeurs perçoivent les 17 approches d'évaluation des apprentissages selon diverses perspectives pédagogiques. Les réponses à ces questions pourront servir notamment à prioriser certaines des diverses approches d'évaluation en formation et perfectionnement. Puisque tous les répondants pouvaient choisir parmi les 17 choix de réponses présentées à la question 34, nous présentons les réponses sous forme de palmarès des cinq approches d'évaluation les plus "populaires". Le tableau P4-a (suite), ci-après, présente ces palmarès.

Tableau P4-a: Les approches utilisées pour évaluer les apprentissages, en ordre décroissant d'utilisation. Le pourcentage est constitué du cumul des "toujours" et "souvent" à la Q34: "Utilisez-vous les approches suivantes pour effectuer l'évaluation de l'apprentissage de vos étudiants?" N = 630

Q34 n)	examen écrit	81%
l)	exercices en classe	60%
a)	devoirs ou exercices à la maison	55%
j)	résolution de problèmes	44%
g)	travail écrit, recherche	42%
k)	démonstration de savoir-faire	40%
f)	rapport ou compte-rendu d'activités ou de laboratoire	36%
h)	questions orales	26%
e)	mini-test ou jeux questionnaire	22%
c)	étude de cas	22%
b)	dossier anecdotique, journal de bord	17%
i)	projet thématique	15%
d)	observation directe avec grille d'appréciation et/ou liste de vérification (check-list)	14%
o)	l'autoévaluation	12%
q)	schéma ou réseau de concepts	9%
p)	l'évaluation par les pairs	4%
m)	examen surprise	3%

Les choix d'instrumentation

La question 42 nous révèle que, pour l'ensemble des répondants, c'est d'abord le professeur seul qui choisit les types d'instruments d'évaluation (44%). C'est le cas surtout en Techniques humaines et en Lettres.

Mais près de 40% des répondants disent qu'ils choisissent les types d'instruments d'évaluation avec leurs collègues donnant le même cours et ce, surtout, en Techniques biologiques.

Ces deux modes de planification sont partagés à peu près également en Sciences biologiques et en Administration (pré-universitaire et technique).

La collaboration avec le département et avec les étudiants est beaucoup plus rare. Les professeurs déterminent leurs évaluation avec le département dans une proportion de 14% seulement (cette proportion est de 22%, un professeur sur cinq, en Techniques physiques). L'implication des étudiants relève de cas isolés (3%).

Tableau P4-a: Palmarès des cinq (5) approches d'évaluation les plus invoquées, pour chacune des questions 35 à 40 (...à partir des fréquences pondérées)

	Q35) Me permet le mieux de savoir si les étudiants ont atteint les objectifs visés	Q36) M'informe le mieux des progrès de l'étudiant	Q37) J'utiliserais davantage, si je la connaissais mieux...
1	examen écrit	examen écrit	schéma ou réseau de concepts
2	travail écrit, recherche	exercices en classe	résolution de problèmes
3	démonstration de savoir-faire	devoirs ou exercices à la maison	autoévaluation
4	résolution de problèmes	travail écrit, recherche	dossier anecdotique, journal de bord
5	devoirs ou exercices à la maison	démonstration de savoir-faire	observation avec grille d'appréciation ou - checklist

	Q38) J'utiliserais davantage, si j'avais plus de temps	Q39) Utile comme aide à l'apprentissage	Q40) Utile pour accumuler des notes au bulletin
1	démonstration de savoir-faire	exercices en classe	examen écrit
2	projets thématiques	résolution de problèmes	exercices en classe
3	exercices en classe	travail écrit, recherche	mini-test ou jeux questionnaires
4	travail écrit, recherche	devoirs ou exercices à la maison	travail écrit, recherche
5	étude de cas	démonstration de savoir-faire	devoirs ou exercices à la maison

Tableau P4-a: (suite) Palmarès des cinq (5) approches d'évaluation formative les plus invoquées à la question 41 (plusieurs réponses étaient possibles)

	Les approches indiquées en 1er choix	Les cinq (5) principales parmi toutes les réponses
1	devoirs ou exercices à la maison	exercices en classe
2	exercices en classe	questions orales
3	questions orales	devoirs ou exercices à la maison
4	journal de bord	résolution de problèmes
5	démonstration de savoir-faire	démonstration de savoir-faire

5.2.5 Cinquième catégorie de Stiggins: Les dimensions interpersonnelles de l'évaluation

Selon Stiggins, les variables interpersonnelles ont un impact sur l'évaluation en salle de classe ainsi que sur l'interprétabilité des notes. Certaines contraintes interpersonnelles viennent des étudiants eux-mêmes, d'autres des enseignants. Par exemple, les attentes et les attitudes envers les individus; la conception de leur rôle dans l'enseignement et dans l'apprentissage, l'évaluation et le contrôle de la salle de classe; le tempérament; la motivation à enseigner et les aptitudes à réagir aux personnes. L'enseignant compétent doit savoir reconnaître et contrôler ces nombreuses dimensions personnelles s'il veut assurer à ses étudiants une évaluation équitable.

Les croyances relatives aux dimensions interpersonnelles de l'évaluation

Les variables "étudiantes"

Les trois premières questions concernent des variables "étudiantes": la motivation et l'effort. Un peu plus de la moitié des répondants (52%) croient que les étudiants qui travaillent fort méritent d'avoir des notes plus élevées que ceux qui ne le font pas (Q44a) (tableau C5-a). Sur ce point, il y a différence significative entre les strates, révélant que c'est en Éducation physique, d'une part, et en Techniques biologiques et en Sciences humaines (incluant Philosophie), d'autre part, qu'on trouve, respectivement, les plus fortes proportions d'accord et de désaccord sur cet énoncé (tableau C5-b, en annexe 5).

Par ailleurs, sept professeurs sur 10 croient que de mauvais résultats d'évaluation ont un effet démotivant sur les étudiants (Q44b). Dans la même proportion, les professeurs répondent qu'il faut s'abstenir d'augmenter la note d'un étudiant pour l'encourager ou le pousser à travailler davantage (Q44c). On constate également que plus d'un professeur sur cinq n'est pas d'accord avec les deux énoncés portant sur

l'effet de la note sur la motivation. Pour ces trois questions relatives à la motivation et à l'effort, près d'un professeur sur vingt (5%) "ne sait pas".

Les variables "professeurs"

Les deux dernières questions de cette série mesurent des variables "professeurs". La question 44d mesure la perception qu'on a de la composante "subjectivité" lors des corrections. Dans l'ensemble, l'opinion est partagée à peu près également entre l'accord et le désaccord, à l'effet que, lors de la correction des questions à développement ou des travaux écrits, les enseignants sont influencés par la connaissance de l'identité du répondant (Q44d) (tableau C5-a). On constate, par ailleurs, que 12% des répondants ne savent que répondre sur ce point.

Tableau C5-a: Pourcentages des réponses aux énoncés de croyances portant sur les dimensions interpersonnelles de l'évaluation

		TA	A	D	TD	NSP
Q44a)	Les plus travaillants méritent notes plus élevées	16%	36%	39%	4%	5%
Q44b)	Les mauvais résultats sont démotivants	9%	61%	23%	1%	5%
Q44c)	Faut pas augmenter une note pour encourager à travailler davantage	15%	55%	21%	3%	6%
Q44d)	Dans la correction, on est influencé par l'identité du répondant	4%	40%	35%	9%	12%
Q44e)	Conception de l'enseignement influence l'évaluation	26%	57%	8%	1%	7%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération des strates. N min = 620; N max = 627

La dernière question de cette série mesure la perception de la relation entre les pratiques d'évaluation et la conception de l'enseignement. Seul un professeur sur dix ne croit pas que "c'est la conception qu'on se fait de l'enseignement qui influence le plus nos choix de stratégies d'évaluation" (Q44e) alors que 7% ne savent pas. Il y a différences statistiques entre les strates sur ces deux dernières questions. Le tableau C5-b, en annexe 5, illustre ces différences.

Les pratiques relatives aux dimensions interpersonnelles de l'évaluation

Les ajustements de groupe

Les questions 23 et 48 portent sur les ajustements que des professeurs offrent à l'ensemble de la classe. Le tableau P5-a montre que, si un professeur sur quatre ne donne jamais un examen plutôt facile pour compenser un examen antérieur difficile (Q23), 37% pratiquent ce genre de compensation parfois ou souvent. Puisqu'il y a différence significative entre les strates sur cette question, le tableau P5-b, en annexe 5, montre que cette pratique est la plus fréquente en Sciences humaines et en Techniques physiques. On remarque également qu'a posteriori, lorsqu'une moyenne générale s'avère plus faible qu'escomptée, un professeur sur dix peut choisir, souvent ou parfois, de modifier les notes de tous ses étudiants pour ajuster la moyenne de la classe (Q48).

Les ajustements individuels

Si on peut ajuster les notes de tout un groupe d'étudiants, on peut aussi choisir d'intervenir sur la note d'un seul individu. Les questions 45, 46 et 47 mesurent cette pratique.

La réponse à la question 47 nous révèle que la grande majorité des professeurs ont tendance à encourager un étudiant faible en accordant une note de passage lorsque ses efforts lui semblent appréciables. Le tableau P5-a montre que près d'un professeur sur cinq le fait au moins "souvent" alors que plus du tiers le fait "parfois". Il y a différence significative entre les strates et le tableau P5 b, en annexe 5, montre que cette pratique est la plus fréquente en Éducation physique et la plus rare en Sciences et Techniques biologiques.

Les questions 45 et 46 demandent aux répondants s'il leur arrive d'intervenir sur des notes finales transmises au bulletin afin de pénaliser des individus pour certains comportements ou pour en récompenser pour d'autres motifs. Le tableau P5-a montre que les professeurs ont plus tendance à augmenter la note qu'à la diminuer. Il y a autant de professeurs qui pénalisent des étudiants pour leurs absences non motivées qu'il y a de professeurs qui récompensent des présences assidues: 18%. Mais si 14% des répondants disent diminuer des notes pour le peu d'effort fourni par certains étudiants, 41% augmentent la note finale en fonction de la motivation et de l'effort dans l'apprentissage. Le tableau P5-a montre les proportions de répondants qui augmentent ou diminuent des notes individuelles en fonction de divers facteurs. Le tableau P5-a est présenté dans l'ordre décroissant des réponses observées.

Tableau P5-a: Pratiques d'intervention sur les notes. "Il m'arrive de..."

Q45)	...Diminuer une note finale pour "pénaliser": N = 618	a)	absences non motivées	14%
		b)	peu d'effort fourni	14%
Q46)	...Augmenter une note finale pour "récompenser": N = 612	e)	progrès réalisés	42%
		a)	motivation et effort	41%
		g)	créativité, originalité	33%
		f)	participation en classe	31%
		d)	réalisation de travaux supplémentaires	23%
		c)	devoirs à la maison	21%
		b)	présence assidue	18%

Tableau P5-a: (suite) Pratiques d'intervention sur les notes. "Il m'arrive de..."

		T	S	P	R	J
Notes de groupes						
Q23)	Donner un examen plutôt facile pour compenser un examen antérieur difficile		5%	32%	37%	26%
Q48)	Modifier les notes de tous pour ajuster moyenne de classe		2%	8%	22%	68%
Notes d'individus						
Q47)	Donner une note de passage à l'étudiant faible si ses efforts sont appréciables	2%	17%	36%	31%	14%

Note: % calculés sur fréquences pondérées.

Le profil de réponses n'est pas différent, entre les strates, pour la réponse (46c) "l'exécution des devoirs ou des travaux à la maison ou en classe". Pour tous les autres facteurs, l'analyse de variance (ANOVA) nous révèle qu'il y a différence significative entre les strates et le tableau P5-b, en annexe 5, illustre les tendances des strates en comparaison avec la moyenne observée dans l'échantillon total. On constate que c'est surtout en Éducation physique et en Arts que les professeurs augmentent ou diminuent des notes finales individuelles et ce, pour la plupart des facteurs mesurés. Cependant, les professeurs de Techniques physiques et de Lettres sont les plus nombreux, proportionnellement, à augmenter des notes pour la réalisation de travaux supplémentaires (Q46d).

5.2.6 Sixième catégorie de Stiggins: La rétroaction et la notation

Les enseignants fournissent continuellement des rétroactions à partir des résultats des évaluations. Selon Stiggins, les professeurs doivent savoir attribuer pertinemment les notes afin que cette forme de rétroaction soit interprétée de façon univoque. Il faut aussi que, dans toutes ses formes, la rétroaction corresponde à ce qui est prévu explicitement dans les objectifs à atteindre, qu'elle soit transmise "à temps" et qu'elle soit signifiante pour l'étudiant.

Les croyances relatives à la rétroaction et à la notation

La rétroaction

Les questions 49e et 49i portent sur certaines dimensions de la rétroaction. L'immense majorité des répondants (90%) croient que les enseignants devraient généralement commenter par écrit les productions écrites de leurs étudiants (Q49e) et 85% disent que la plupart des étudiants lisent les commentaires qu'ils écrivent sur les feuilles-réponses ou sur les travaux écrits (Q49i) (tableau C6-a). Seuls 7% des professeurs ne croient pas que la plupart des étudiants lisent les commentaires que leurs professeurs écrivent sur leurs feuilles-réponses ou sur leurs travaux écrits. Par ailleurs, on note qu'il y a différence significative entre les strates sur la question 49e, révélant que les plus fortes proportions de désaccord sont dans les strates Lettres et Techniques physiques (tableau C6-b).

Tableau C6-a: Pourcentages des réponses aux énoncés de croyances portant sur la rétroaction et la notation

		TA	A	D	TD	NSP
Rétroaction						
Q49e)	Les professeurs devraient commenter par écrit les productions écrites des étudiants	33%	57%	4%	1%	5%
Q49i)	La plupart des étudiants lisent les commentaires écrits par leurs professeurs	25%	60%	6%	1%	9%
Notation						
Q49a)	La moyenne est le reflet de la qualité de l'enseignement	1%	17%	62%	14%	6%
Q49f)	Faut éviter les évaluations qui impliquent la subjectivité	18%	44%	25%	4%	8%
Q49d)	Des professeurs accordent des notes de passage non méritées	4%	28%	23%	4%	41%
Q49b)	Les notes devraient suivre la courbe normale	2%	38%	39%	12%	10%
Q49g)	L'évaluation formative ne doit pas compter au bulletin	16%	35%	32%	7%	10%
Q49c)	La notation nuit à l'enseignement	2%	9%	57%	24%	8%
Q49h)	La note, c'est la paie de l'étudiant	9%	46%	26%	9%	10%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération des strates. N min = 622; N max = 628

La notation

Tous les autres énoncés de cette série portent sur diverses croyances relatives à l'attribution des notes qui sont éventuellement versées au bulletin.

Trois énoncés, 49a, 49f et 49d, portent sur la crédibilité des notes. Le tableau C6-a montre que la très grande majorité des répondants ne croient pas que la note moyenne de la classe est un reflet direct de la qualité de

l'enseignement reçu (Q49a) alors que 18% des répondants signifie leur accord à cet énoncé.

De plus, pour l'ensemble des strates, 63% des répondants croient qu'il faut éviter de faire des évaluations qui impliquent le jugement personnel, la subjectivité de l'enseignant (Q49f). Pourtant, un professeur sur quatre (25%) est en désaccord et quatre professeurs sur 100 sont très en désaccord, affirmant implicitement que le jugement personnel et une certaine subjectivité peuvent avoir une place légitime en évaluation. Notons ici, que 8% des répondants ne savent que répondre à cette question (tableau C6-a).

Le troisième énoncé qui mesure les croyances relatives à la crédibilité des notes est celui qui demande aux répondants de porter un jugement de valeur sur l'hypothèse que plusieurs enseignants accordent une note de passage à des étudiants qui ne la méritent pas (Q49d). Cet énoncé a récolté un nombre très élevé de "ne sais pas" (tableau C6-a), probablement à cause des incidences déontologiques évidentes de cette question. Quant aux autres répondants, 32% croient que plusieurs enseignants accordent une note de passage à des étudiants qui ne la méritent pas alors que 27% sont en désaccord avec cet énoncé. Le tableau C6-a montre qu'il y a une quasi-symétrie entre les accords et les désaccords à cet énoncé: 4% très en accord et 4% très en désaccord; 28% en accord et 23% en désaccord. Il y a différence significative entre les strates sur cet énoncé et le tableau C6-b, en annexe 5, montre les maxima et minima.

Deux autre énoncés mesurent deux aspects de la conception qu'on peut se faire de la genèse des notes. D'abord, l'énoncé 49b concerne le concept de distribution des notes selon la courbe normale de probabilités. De l'ensemble des répondants, 40% croient que, dans un groupe, la distribution des notes doit suivre la courbe normale, en forme de cloche: peu de faibles, beaucoup de moyens et peu de forts. Par contre, 51% des professeurs sont en désaccord, parfois même très en désaccord (12%) avec ce concept. Les proportions sont inversées à propos de la participation de l'évaluation formative à la genèse des notes finales. En effet, 51% croient que les résultats d'une évaluation formative ne doivent pas servir au

calcul de la note inscrite au bulletin (Q49g) et 39% sont en désaccord (dont 7% très en désaccord). Sur cette question, il y a différence significative entre les strates et le tableau C6-b, en annexe 5, montre que, si les accords se trouvent notamment en Sciences biologiques, en Sciences physiques et en Techniques humaines, les désaccords les plus nombreux, toutes proportions gardées, sont en Arts (pré-universitaire et technique) et en Techniques physiques. Remarquons que, sur ces deux énoncés, un professeur sur dix ne sait que répondre.

La question 49c mesure une croyance relative à l'effet de la notation sur le processus d'enseignement. On apprend que la grande majorité des répondants (81%) ne croient pas que la notation nuit à l'enseignement. Mais un professeur sur dix croit que la notation est nuisible et 2% affirment même être très en accord avec l'énoncé. Quant à l'énoncé 49h, on voit que la majorité (55%) des professeurs croient que la note, c'est la paie de l'étudiant.

Les pratiques relatives à la rétroaction et à la notation

La notation

La question 28 nous apprend que la grande majorité des professeurs, lorsqu'ils évaluent un travail fait par une équipe d'étudiants, accordent la même note à tous ses membres. En effet, lorsque nous éliminons les professeurs qui ont répondu que cette question ne s'applique pas à leur situation (22%), on constate que pour ceux qui sont concernés, 40% accordent toujours la même note à tous les membres de l'équipe et 40% accordent "souvent" la même note. Le tableau P6-b, en annexe, ilustre les maxima et minima par chacune des strates.

Dans le calcul de la note finale qui est inscrite au bulletin de l'étudiant, deux répondants sur cinq (42%) n'incluent jamais des notes obtenues dans des évaluations formatives (Q50), 58% le font au moins rarement, et 15%, toujours. C'est en Sciences humaines (incluant Philosophie) et en

Arts (pré-universitaire et technique) qu'on retrouve les plus fortes proportions de "toujours" (tableau P6-b).

Tableau P6-a: Description des pratiques de rétroaction et de notation

		T	S	P	R	J
Q28)	Pour un travail fait par une équipe d'étudiants, j'accorde la même note à tous les membres de l'équipe	33%	29%	9%	4%	3%
Q50)	La note finale contient des notes obtenues en évaluation formative	15%	17%	12%	13%	42%
Q52)	Après un examen ou un travail, je donne à chacun une rétroaction détaillée sur les points à améliorer	23%	44%	24%	7%	2%
Q53)	Après un examen ou un travail, je prends des moyens pour que les étudiants puissent améliorer leurs points faibles	23%	56%	18%	3%	----

La rétroaction

Lorsqu'ils retournent un examen ou un travail aux étudiants, les professeurs donnent à chacun une rétroaction détaillée sur les points à améliorer (Q52) et ce, plutôt souvent (44%) que toujours (24%). De même, un professeur sur quatre ne fait cette rétroaction détaillée que parfois. Le tableau P6-b montre que c'est surtout en Sciences biologiques, en Éducation physique et en Administration (pré-universitaire et technique) que les proportions de "rarement" dépassent les 10%.

Enfin (Q54), les examens et les travaux, une fois corrigés, sont généralement retournés aux étudiants au plus tard après deux semaines (96%), la très grande majorité les retournant après une semaine (72%). Dans quelques strates seulement, certains professeurs mettent trois semaines pour retourner aux étudiants les examens et travaux corrigés: 9% en Techniques humaines, 7% en Administration et 11% en Français.

La remédiation

Après un examen ou un travail, 79% des répondants prennent souvent ou toujours (23% disent toujours) les moyens nécessaires pour que les étudiants puissent améliorer leurs points faibles (Q53). Mais près d'un professeur sur cinq (18%) prend ces moyens "parfois". Si personne ne prend jamais de tels moyens, 3% de l'ensemble des répondants le font rarement et ce, surtout en Sciences biologiques (9% de cette strate) et en Administration (6% de cette strate).

La politique d'évaluation

Dans la détermination des notes finales (Q51), la plupart tiennent compte de la politique d'évaluation des apprentissages adoptée par leur département ou par leur collège (64% toujours et 25% souvent). Seuls quelques cas isolés n'en tiennent compte que rarement sinon jamais (5%).

5.2.7 Les questions portant sur la formation et le perfectionnement

La fin de notre questionnaire est composée de quatre questions destinées à mesurer certaines dimensions relatives à la formation des professeurs en évaluation ainsi qu'à la perception des besoins de perfectionnement. De plus, une section "Informations générales" recueille des données de type démographique afin de mieux connaître les répondants et de les associer à nos strates.

Formation

La question 55 nous informe du niveau de préparation académique des enseignants dans le domaine de l'évaluation des apprentissages. Un choix de réponses permet à certains répondants de signifier qu'ils n'ont pas de formation, formelle ou informelle, en évaluation.

Dans le tableau Q55, ci-après, les types de préparation sont placées en ordre décroissant, selon les proportions calculées sur les fréquences pondérées. On voit que les professeurs de cégeps ont d'abord reçu de la formation en évaluation dans leur établissement d'enseignement (journée pédagogique, cours du CPEC-Performa, conférence ou atelier) ou encore, se sont donnés une formation personnelle, en autodidactes. La formation universitaire, au baccalauréat ou à la maîtrise, est au dernier rang des sources de formation. Enfin, 14% des répondants disent ne pas avoir reçu de préparation en évaluation des apprentissages.

Q55)	"Quelle préparation avez-vous reçue en matière d'évaluation des apprentissages?"	
Rang		
1-	Au moins une journée pédagogique	40%
2-	Autodidaxie	39%
3-	Au moins un cours au CPEC-Performa	37%
4-	Au moins une conférence ou un atelier	35%
5-	Au moins un cours au baccalauréat	32%
6-	Au moins un cours à la maîtrise	15%
7-	Aucune	14%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération. N = 628

Si on exclut les journées pédagogiques et l'autodidaxie pour lesquelles les réponses des professeurs ne varient pas significativement d'une strate à l'autre, il y a différence entre les strates pour toutes les autres réponses de la question 55. Le tableau Q55-b, en annexe 5, illustre les maxima. On y notera que c'est en Sciences humaines qu'on trouve la plus forte proportion de répondants ayant indiqué qu'ils n'ont reçu aucune préparation antérieure en évaluation.

Nous avons ensuite demandé aux répondants d'autoévaluer leurs compétences en évaluation formative et sommative (Q56). La majorité des professeurs jugent que leur compétence est plutôt bonne, sinon très bonne en évaluation formative (67%), de même qu'en évaluation sommative (88%). Le tableau Q56, en annexe 5, montre que très peu d'enseignants jugent que leur compétence est faible. À l'égard de l'évaluation sommative, personne ne juge nulle sa compétence. Ici encore, il y a différences significatives inter-strates et le tableau Q56, en annexe 5, montre clairement que les professeurs sont plus à l'aise avec le concept d'évaluation sommative qu'avec l'évaluation formative. Tenant compte qu'à la question précédente les répondants de la strate Sciences humaines (incluant Philosophie) disaient être les plus nombreux à n'avoir aucune formation antérieure en mesure et évaluation (24%), on ne sera pas surpris de voir que c'est cette même strate qui

participe aux plus fortes proportions dans la colonne "compétence faible ou nulle en évaluation formative".

Perfectionnement

Dans la perspective d'atteindre partiellement notre objectif d'identifier des pistes de perfectionnement à offrir aux enseignants des cégeps aux lendemains de la présente recherche, les questions 57 et 58 proposaient quelques choix de cibles de perfectionnement et de ressources.

La question 57 nous apprend que les cibles prioritaires de perfectionnement seraient les suivantes, dans l'ordre décroissant des réponses obtenues:

Q57)	"Si on organisait des activités de perfectionnement en mesure et évaluation, quels sujets, parmi les suivants, vous intéresseraient davantage? (vous pouvez cocher plusieurs sujets)"	
1-	Évaluation fondée sur les compétences à acquérir	48%
2-	Évaluation formative	47%
3-	Évaluation des attitudes	40%
4-	Construction d'examens	37%
5-	La correction (annotation)	25%
6-	Techniques d'observation	23%
7-	Planification de l'évaluation	21%
8-	Rédaction d'objectifs	19%
9-	L'attribution des notes	16%
10-	Aucune	10%
11-	Autre (précisez)	6%

Note: Ces % proviennent des tableaux de fréquences calculées après pondération. N = 624

Un professeur sur dix indique qu'aucune activité de perfectionnement ne l'intéresserait. Par ailleurs, 6% des répondants ajoutent des sujets que nous n'avions pas inclus dans nos choix de réponses. Ces sujets portent:

- sur les cibles de l'évaluation:
  - les laboratoires et ateliers;
  - les productions écrites;
  - les travaux de création;
- sur l'instrumentation:
  - la coévaluation et l'autoévaluation;
  - la rédaction d'examens à questions à choix multiples;
  - comment diversifier les dispositifs d'évaluation;
- sur les qualités de l'évaluation:
  - la validation des questions à choix multiples;
  - l'analyse des items;
- sur la notation:
  - la "correction" des études de cas.

La question 58 a permis aux répondants d'indiquer un certain ordre de priorités parmi un choix de ressources qui pourraient leur être utiles.

Dans l'ordre décroissant, les répondants indiquent les besoins suivants (les proportions sont fondées sur les fréquences pondérées):

Q58)	"Quelles ressources, parmi les suivantes, vous seraient utiles?"	
1-	Des examens modèles	36%
2-	Des modèles de consignes pour travaux écrits	34%
3-	Une banque de questions d'examens	30%
4-	Un système informatisé d'analyse d'examens	26%
5-	Aucune	23%
6-	Autre (précisez)	11%
7-	Un système informatisé de correction d'examens à choix multiples	10%
8-	Des examens de fin de session déjà construits	9%

Note: Proportions calculées sur fréquences pondérées. N = 618

Si 23% ont coché le choix "aucune", à l'effet qu'ils ne souhaitent aucune des ressources proposées, 11% ont ajouté d'autres ressources qui ne faisaient pas partie de nos propositions. Les plus signifiantes, demandées par cinq répondants et plus, sont:

- de l'épaulement:
  - de la formation individualisée, du counselling en évaluation;
  - de la formation de groupe (en département): séminaires et conférences;
- des outils:
  - des grilles d'évaluation et de correction;
  - des exemples, des modèles d'instruments de mesure;
- des aménagements administratifs:
  - réduire la tâche et avoir du temps;
  - accorder des ressources humaines, notamment des assistants de correction d'examens ou de travaux.

Parmi les ressources qui sont souhaitées par les répondants, on remarque que, les réponses nous permettent de dégager au moins trois profils de ressources.

D'abord, les professeurs souhaitent qu'on leur propose des modèles qui pourront les aider à résoudre des problèmes pratiques en évaluation: modèles d'examens et modèles de consignes de travaux écrits. Dans les réponses manuscrites, plusieurs répondants ajoutent à ce chapitre. On souhaite avoir des exemples concrets, des trucs, des instruments prérédigés qui contribueront au savoir-faire des professeurs à l'égard des divers objets ou cibles à évaluer dans le cadre de leur tâche.

Ensuite, les répondants souhaitent de l'aide informatisée: des banques d'items, des systèmes de correction et d'analyse. Ce profil de ressources a manifestement deux attraits. L'aide informatisée peut contribuer à améliorer la qualité de l'évaluation tout en réduisant le fardeau de la tâche.

Enfin, les professeurs souhaitent que des ressources humaines soient mises à contribution, soit pour alléger la tâche de correction des travaux, soit pour les aider, individuellement ou en groupe, à parfaire leur compétence en évaluation ou à résoudre des problèmes ponctuels. Sur ce dernier point, plusieurs répondants ont précisé que des séminaires ou conférences départementales, des lectures guidées et du counselling individuel en mesure et évaluation constitueraient des ressources utiles.

Bien entendu, des professeurs ont dit souhaiter que la tâche d'enseignement soit réduite afin de disposer de plus de temps, notamment pour effectuer l'évaluation des apprentissages de leurs étudiants.

Ainsi, les ressources demandées par les professeurs s'inscrivent sous la triple préoccupation de résoudre des problèmes, de réduire le fardeau de la tâche et d'accroître leur compétence.

Informations générales

La dernière section du questionnaire recueille des renseignements qualifiés de "démographiques" au sujet des répondants. Enfin, une question cherche à connaître l'arrière-plan pédagogique de nos répondants.

Les répondants appartiennent aux 47 collèges et campus qui ont été identifiés lors de la constitution de l'échantillon. Le lecteur est invité à se reporter au chapitre 4 pour y lire la description de l'échantillon prévu ainsi que les taux de réponses obtenues dans chaque établissement. Ce même chapitre décrit aussi la représentativité des strates ainsi que les poids qui ont été attribués à chaque strate pour corriger des déséquilibres qui se sont manifestés suite aux retours des questionnaires. Les proportions de femmes et d'hommes sont, respectivement, 37% et 63%.

Nos répondants ont un âge moyen de 43 ans. Le plus jeune a 23 ans et le plus vieux, 71 ans.

En moyenne, ils ont 16 années d'expérience en enseignement et on leur reconnaît 18 années de scolarité pour fins de rémunération.

La question 65 décrit le niveau de scolarité des professeurs qui ont répondu au questionnaire.

Tableau Q65: Niveaux de scolarité des répondants, par ordre décroissant. (N = 622)

Rang	Scolarité	Proportion de répondants
1	Baccalauréat	47%
2	Maîtrise	38%
3	Doctorat	8%
4	Certificat	4%
5	DEC	3%

18% des répondants ont ajouté qu'ils possèdent aussi une scolarité dans un autre champ que la spécialité enseignée.

Nous avions choisi de proposer à nos répondants de nous indiquer les formules pédagogiques qu'ils utilisent (Q66). Chacune des strates a donné lieu à des profils très variés. Puisqu'il y a différence significative entre les strates, le tableau 66, en annexe, montre les maxima et les minima pour chaque formule pédagogique proposée. Sur plusieurs points, ce tableau n'offre pas de surprise: il fallait s'attendre à ce que le travail en équipe soit important en Éducation physique, le laboratoire le soit en Sciences biologiques et le stage en Techniques biologiques. Mais on retrouve aussi des indications fort intéressantes sur l'usage que font les diverses strates de certaines approches comme le séminaire (ou groupe de discussion), le travail en équipe, le laboratoire. Le taux de réponses obtenues à la formule "exposé informel en classe" peut étonner. Cette formule se place au cinquième rang (tableau Q66) alors que l'exposé magistral formel se place au premier rang. Il y aurait lieu de se demander comment les professeurs interrogés ont compris le libellé de ces deux formules pédagogiques.

CHAPITRE 6
ANALYSE DES RÉSULTATS

Dans le chapitre précédent, nous avons présenté les résultats obtenus par notre questionnaire auprès des répondants. En ce sens, le chapitre de présentation des résultats a permis d'atteindre nos deux premiers objectifs, à savoir:

1. Identifier les croyances des enseignants des cégeps à l'égard de l'évaluation des apprentissages en salle de classe;

2. Décrire les pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages et ce, en situation régulière de salle de classe;

Le présent chapitre est consacré à l'analyse des résultats et, de ce fait, nous permettra d'atteindre notre troisième objectif de recherche:

3. Dégager des pistes de perfectionnement qui prennent en compte les croyances et les pratiques des enseignants du collégial en matière d'évaluation des apprentissages en salle de classe.

6.1 Considérations d'analyse

Dans ce qui suit, nous analysons, comparons et commentons les données à partir de leur appartenance à chacun des six champs de compétence en évaluation, tels que proposés par Stiggins.

Pour chaque champ de compétence (ou catégorie), les croyances et les pratiques sont regroupées de façon à mettre en lumière, s'il y a lieu, des liens, des zones de cohérence ou de contradiction entre les croyances d'une part et les pratiques d'autre part.

Dans les discussions qui suivent, nous faisons intervenir les variables dépendantes et indépendantes issues du questionnaire, ainsi que des données qualitatives. Les variables dépendantes correspondent aux questions 1 à 54 du questionnaire et les variables indépendantes sont celles qui font l'objet des questions 55 à 66. Ces dernières comportent des données

démographiques (discipline enseignée, âge, sexe, années d'expérience, scolarité) et des données relatives à la formation et aux besoins en évaluation. Les données qualitatives sont constituées des commentaires formulés par les groupes réacteurs.

6.1.1 Les groupes réacteurs: Attitudes générales

Nous avons présenté les résultats préliminaires de la recherche à trois groupes hétérogènes de professeurs. Ces groupes de professeurs constituent les groupes réacteurs décrits au chapitre 4 (la méthodologie). Tel que recommandé par Guba et Lincoln (1991, p. 110) ainsi que par Huberman et Miles (1991, p. 442), ce retour à la source a pour but de tester la crédibilité, la plausibilité de nos données et ce, auprès des informateurs eux-mêmes. Essentiellement, nous cherchons à confirmer les résultats du questionnaire. Ces trois groupes de professeurs à qui nous avons présenté les mêmes résultats statistiques ont fait des commentaires qu'on peut regrouper en cinq catégories:

1. Crédibilité

Personne n'a soulevé de questions ou émis de commentaires tendant à mettre en doute la validité de nos résultats. Plusieurs nous ont dit y voir confirmation de pratiques incorrectes et ont fait des commentaires du type: "il faut corriger cette situation; il est temps qu'on s'en occupe".

2. Justification

Certains tendent à interpréter les chiffres en apportant des explications circonstancielles qui justifieraient les pratiques observées: "Cette pratique est normale, moi aussi je fais ça"; "Dans la réalité de la salle de classe, on est obligé de faire ça, pour motiver les étudiants, pour les faire travailler"; "La théorie, c'est une chose, la réalité, c'est autre chose".

3. Inquiétude

Certains s'inquiètent de ce qu'on va faire avec ces données, maintenant qu'on a chiffré les pratiques des professeurs: "Allez-vous recommander que des normes de pratique soient imposées aux professeurs?" L'insistance, voire l'agressivité de quelques-uns, nous fait observer l'émergence d'une réaction que nous n'avions pas prévue: la peur.

4. Priorité aux pratiques

Les commentaires concernaient presque tous des énoncés de pratiques. Lorsque des énoncés de croyances étaient évoqués, c'était pour soulever des contradictions apparentes dans les réponses à des énoncés apparentés de croyances et de pratiques.

5. Classement en pratiques correctes ou incorrectes

Tous mettent l'accent sur les chiffres qui décrivent des pratiques qui sont généralement perçues comme incorrectes. On a tendance à lire les résultats avec une grille subjective qui classe les pratiques selon qu'elles sont correctes ou incorrectes, bonnes ou mauvaises. Comme si chaque échelle de réponses à nos items constituait une échelle d'évaluation graduée, du plus correct au moins correct. En lisant les tableaux de proportions observées, les gens présents ont tendance à réagir à des données extrêmes, par exemple, lorsque la proportion de "jamais" est forte.

6.1.2 Les sources de variation dans les profils de pratiques

Plusieurs des commentaires reçus des groupes réacteurs nous amènent à reconnaître des facteurs qui pourraient expliquer des variations dans les profils de pratiques, tels que nous les observons. En effet, dans le milieu enseignant dans son ensemble, comme à l'intérieur de chacune de nos

strates, plusieurs facteurs peuvent venir moduler les pratiques observées en évaluation des apprentissages:

1. La tâche d'enseignement, avec ses innombrables variantes caractéristiques de la diversité des disciplines enseignées, constitue une première source de pression sur les pratiques des professeurs en matière d'évaluation.

2. Dans chaque discipline, les pratiques pédagogiques varient beaucoup. Les professeurs choisissent des pratiques d'évaluation qui conviennent aux réalités pédagogiques propres à chaque discipline. Nous considérons que les réponses obtenues à notre questionnaire reflètent ce qui se fait dans le réseau collégial et, notamment, ce qui se fait à l'intérieur des strates. Les écarts de proportions entre les strates, au chapitre 5, sont le reflet des besoins, des coutumes et des réalités disciplinaires propres aux strates de programmes. On n'évalue pas les apprentissages en arts comme on évalue les apprentissages en sciences humaines. Une production artistique ne s'évalue pas nécessairement comme une production littéraire ou comme une production technique.

3. Le cours enseigné, voire la clientèle, influence les pratiques. Dans notre questionnaire, nous demandons au répondant de fonder toutes ses réponses sur l'un de ses groupes-cours. Le profil de réponses pourra donc varier en fonction du cours-témoin choisi par chaque répondant.

4. La conception de l'enseignement et la conception de l'évaluation sont probablement des construits qui peuvent être mis à contribution pour expliquer les différences de pratiques entre les professeurs, quoique la présente recherche ne fait pas l'étude de ces construits.

5. La formation des professeurs en évaluation des apprentissages constitue aussi une source importante de modulation des pratiques observées. Certains ont une formation académique formelle en matière d'évaluation, d'autres fondent toutes leurs pratiques sur une formation folklorique (Stiggins, 1989) acquise dans leur milieu

de travail, plus particulièrement dans leur département (Robitaille et Maheux, 1993).

6.1.3 Jugements de valeur

Dans l'analyse des résultats, nous tenons compte de ces divers facteurs qui peuvent faire varier les profils des pratiques, tels que nous les observons dans nos données. Il n'est donc pas question pour nous de porter des jugements de valeur qui seraient de nature à évaluer les pratiques ou les croyances des professeurs de cégeps à l'égard de l'évaluation des apprentissages. Nous ne faisons pas ici d'évaluation des croyances et des pratiques car toute évaluation devrait nécessairement être fondée sur une comparaison entre les recommandations d'experts en évaluation d'une part, et les pratiques et croyances observées au collégial, d'autre part. Un tel travail de comparaisons et de jugements de valeur n'est pas l'objet de la présente recherche.

Dans les lignes qui précèdent, nous rappelons les objectifs de la recherche qui sont de décrire et de suggérer des pistes de perfectionnement. Malgré tout, le lecteur nous pardonnera si, en certains endroits, notre texte semble évaluatif. On comprendra aisément que certaines des pratiques, observées chez certains répondants, obligeront à y cibler des efforts de perfectionnement. Le lecteur y verra alors que, si certains jugements de valeur sont sous-jacents, ceux-ci sont fondés sur le gros bon sens et sur l'observation que l'immense majorité des professeurs ont des pratiques conformes à ce gros bon sens.

6.1.4 Structure du chapitre

Dans l'ensemble de ce rapport de recherche, les six catégories de Stiggins, en annexe 7, nous servent de modèle structurant. Au chapitre 5, nous référons à ces six catégories pour séparer les uns des autres nos objets d'enquête. Au présent chapitre, nous évoluons vers le concept de compétence parce que nous cherchons à guider la formation dans des champs

spécifiques de compétence. C'est pourquoi nos six catégories sont désignées, ci-après, comme six champs de compétence.

Nous faisons donc six analyses, une pour chacun des champs de compétence proposés par Stiggins (1991). Pour chacune des six études qui suivent, fondées sur les six champs de compétence inspirés de Stiggins (1991), ainsi que pour la septième étude portant sur la problématique plus générale de la formation en mesure et évaluation, nous faisons intervenir:

- les résultats décrits au chapitre 5 et;
- les commentaires des groupes réacteurs.

6.1.5 Analyses statistiques

La présente recherche est descriptive et cherche à décrire ce qui se passe "sur le terrain". Cependant, dans l'esprit de notre troisième objectif de recherche, nous choisissons d'explorer les regroupements d'énoncés qui peuvent être dégagés des profils de réponses obtenues. Pour cette raison, nous utilisons, à l'instar de Frary (1992), des analyses factorielles. Cela nous permet d'identifier des structures et, si possible, de nommer ces structures.

L'analyse factorielle constitue une analyse multivariée qui, espérons-le, guidera notre démarche qui consiste à proposer des pistes de perfectionnement en mesure et évaluation. Cette procédure statistique permet de voir des associations qui peuvent être faites entre des croyances et des pratiques. Même s'il n'est pas dans le mandat de la présente recherche d'étudier les structures causales qui peuvent exister entre des croyances et des pratiques, non plus que la direction des liens de causalité, nous croyons qu'une simplification de nos variables, par analyse factorielle, peut contribuer à une synthèse des données en proposant des profils, parfois caricaturaux, des croyances et des pratiques.

Les analyses factorielles sont effectuées sur le logiciel SPSS - Windows et nous choisissons d'utiliser la méthode de rotation VARIMAX, car cela est réputé produire des facteurs simples et faciles à interpréter (Fontaine, 1988, pp. 142-143; Tabachnick et Fidell, 1989, p. 628).

À cause du très grand nombre de variables suscitées par notre questionnaire, nous choisissons de faire des analyses factorielles distinctes pour chacun des six champs de compétence de Stiggins (1991; annexe 7). Les facteurs identifiés par ces analyses sont présentés, ci-après, dans les champs de compétence auxquels ils appartiennent. Notons tout de suite que, dans le quatrième champ de compétence, l'analyse factorielle n'est pas signifiante et n'est donc pas reproduite. Une septième analyse factorielle est faite sur un groupe de variables pouvant témoigner des "conceptions" que les professeurs se font de l'évaluation des apprentissages. Pour faire cette dernière analyse factorielle, nous choisissons des variables, sans égard à leur appartenance à l'un ou l'autre champ de compétence, en fonction de la probabilité qu'elles contribuent à décrire des conceptions de l'évaluation. L'analyse factorielle sur les conceptions de l'évaluation est incorporée à la section 6.3, traitant de la formation en mesure et évaluation.

Chacune de ces analyses factorielles est faite sur un nombre réduit de variables. Les énoncés ayant des saturations supérieures à 0,30 sont ceux qui permettent d'identifier et de nommer les facteurs. À cause des nombres restreints de variables, certains facteurs sont associés à aussi peu que deux énoncés. Les tableaux statistiques relatifs à chacune des analyses factorielles sont présentés en annexe 6. Pour chaque analyse factorielle, ces tableaux contiennent:

- la liste des variables incluses et leurs communalités; les "eigenvalues" et les pourcentages de variance expliquée pour chacun des facteurs extraits;
- la matrice de facteurs obtenus après rotation VARIMAX.

En plus des analyses factorielles, nous utilisons, quoique plus rarement, des tableaux croisés (tableaux de contingence) pour isoler certaines variables et pour comparer les réponses reçues sur deux variables en même temps. On s'assure que ces analyses sont statistiquement significatives, à 95% et plus de probabilité de certitude, et cette validation est fondée, à chaque fois, sur l'une ou l'autre procédure suivante, selon les variables en jeu.

- le "V" de Cramer lorsque l'on croise une échelle ordinale avec une échelle nominale;
- le "tau" de Kendall lorsqu'on croise deux échelles ordinales.

6.1.6 À propos des recommandations

Les recommandations sont intégrées à chaque analyse lorsqu'elles sont relatives aux champs de compétence. D'autres recommandations sont proposées en conclusion lorsqu'elles portent sur des dimensions plus générales.

Cette recherche explore un domaine qui a été peu étudié au collégial. Le lecteur ne sera pas surpris qu'en conséquence, nos analyses soulèvent plusieurs autres questions ou pistes de recherches qui, évidemment, devront faire l'objet de travaux futurs.

6.2 Analyse des résultats

6.2.1 Le premier champ de compétence: L'utilisation de l'évaluation en classe

Analyse

L'utilisation de l'évaluation pour gérer la classe

On a vu au chapitre 5 que, pour 81% des répondants, les travaux et examens soumis aux étudiants contribuent "toujours" au calcul de la note inscrite au bulletin. Lorsqu'on isole ces répondants pour étudier les réponses qu'ils donnent à certaines questions du questionnaire, on constate que:

- 83% d'entre eux croient que la meilleure façon de faire travailler les étudiants est de donner des points pour leur travail (Q6a; τ = ,206; p = ,0001)
- 96% croient que le fait d'être évalués encourage les étudiants à fournir plus d'effort à l'étude et au travail scolaire (Q6b; τ = ,178; p = ,0001);
- seulement le tiers d'entre eux (34%) ne croient pas que la note représente la paye de l'étudiant (Q49h; τ = ,121; p = ,0001).

Ainsi, presque tous misent sur l'évaluation sommative, sur l'attribution de notes aux travaux et examens, pour motiver les étudiants à investir dans leur travail scolaire et dans l'étude. Une majorité considère même que cela correspond aussi, en quelque sorte, à une paye pour l'étudiant.

Temps de classe consacré à des activités d'évaluation

Le temps de classe consacré à des activités d'évaluation de toute nature varie beaucoup et, si le tiers des professeurs utilisent 10% et moins du cours à des activités d'évaluation, un autre tiers y consacre 20% et plus du

temps de classe (9% y consacre même 40% et plus du temps de classe). En groupes réacteurs, des professeurs ont réagi à ces chiffres en se demandant si, en consacrant tant de temps à l'évaluation, les enseignants ont encore le temps d'enseigner et les étudiants ont encore le temps d'apprendre. L'impact pédagogique d'utiliser beaucoup de temps à l'évaluation peut s'analyser à partir de la question de savoir si les activités d'évaluation incluent l'évaluation formative et si les pratiques d'évaluation formative de ces professeurs sont de nature à interrompre le processus d'apprentissage (par exemple: un mini-test) ou de nature à coïncider avec des activités d'apprentissage (par exemple: le travail en équipe). On sait, en effet, que la moitié des répondants utilisent "toujours" l'examen écrit. Lorsqu'on ajoute ceux qui s'en servent "souvent", on voit que 81% des professeurs des cégeps utilisent souvent ou toujours l'examen écrit. Pour l'ensemble des approches d'évaluation, c'est le dispositif le plus utilisé. Mais c'est aussi un dispositif qui, par définition, interrompt l'enseignement. Or, lorsqu'on analyse de façon très spécifique les pratiques des professeurs à partir de la question 5 portant sur le temps de classe consacré à l'évaluation, on constate que, chez les 30% de professeurs qui font des activités d'évaluation pendant 20% et plus du temps de classe, un professeur sur deux (49%) utilise "toujours" l'examen écrit. On constate, ici, que le processus d'enseignement-apprentissage semble souvent interrompu dans les classes de certains professeurs et ceux-là pourraient tirer profit à expérimenter l'utilisation au moins partielle de dispositifs de mesure qui coïncident avec les activités d'apprentissage.

Fréquence de l'évaluation sommative

Une analyse par tableau croisé entre les questions 6g et 11 (V de Cramer: ,097; p = ,0368) nous montre que les professeurs font de l'évaluation sommative, le plus fréquemment, aux trois semaines (46%), même si 10% d'entre eux ne croient pas que les étudiants travaillent plus régulièrement lorsque l'évaluation est fréquente (Q6g). Ces professeurs ont donc adopté cette cadence en évaluation sommative pour d'autres raisons. On observe aussi la situation inverse. Un professeur sur 5 (20%)

fait des évaluations sommatives "aux 5 semaines ou plus rarement", mais 85% d'entre eux, une forte majorité, croient que les étudiants travaillent de façon plus régulière lorsque l'évaluation est fréquente (Q6g). L'incohérence n'est peut-être qu'apparente. Peut-être pratiquent-ils, entre leurs évaluations sommatives, des évaluations formatives qui ne "comptent pas au bulletin", conformément à ce qui est recommandé.

Enfin, si 16% des professeurs de cégeps font de l'évaluation sommative de façon hebdomadaire, on note que la proportion s'élève à un professeur sur cinq (20%) dans les strates Sciences biologiques, Arts et Lettres (V de Cramer : ,184; p = ,0001).

Que l'évaluation soit sommative ou formative dans sa fonction, nous avons vérifié si les travaux et les examens soumis aux étudiants contribuent au calcul de la note inscrite au bulletin. Lorsqu'on isole les professeurs qui font à toutes les semaines de l'évaluation sommative, on constate que, dans les strates Sciences humaines et Sciences physiques, tous disent que les travaux et examens contribuent "toujours" au calcul de la note inscrite au bulletin (V de Cramer: ,389; p = ,0231). Dans ces strates, donc, il y a des professeurs qui évaluent les apprentissages à toutes les semaines et ces évaluations comptent toutes pour le bulletin.

Dans toutes les autres strates, une proportion minoritaire de répondants disent que les examens et travaux contribuent "souvent" au calcul de la note finale. Si ces professeurs font de l'évaluation sommative à toutes les semaines, ils disent quand même que ces évaluations contribuent souvent au bulletin, nous laissant devant trois hypothèses: ou bien ils font aussi de l'évaluation formative en plus des évaluations sommatives hebdomadaires, ou ils ont répondu en pensant aux exercices qu'ils prescrivent mais qui ne contribuent pas au calcul de la note, ou, enfin, ils ont répondu de façon incohérente à ces questions.

De tous ceux qui évaluent à des fins sommatives "aux trois semaines" (Q11), la question 50 nous apprend que, 42% d'entre eux ne versent "jamais" au bulletin des notes obtenues dans des évaluations formatives. Seuls 12% versent toujours au bulletin des notes obtenues en évaluation

formative (V de Cramer : ,129; p = ,001). Certains professeurs évaluent si souvent qu'ils n'ont pas le temps de faire des évaluations formatives ou font en sorte que les éventuelles évaluations formatives se confondent à l'évaluation sommative. À l'autre extrême, d'autres évaluent plus rarement en vue du bulletin final mais donnent à croire qu'une minorité d'entre eux pratiquent des évaluations formatives.

Des pratiques qui génèrent des réticences

Cela mène à se demander si ces pratiques d'évaluation sont toutes de nature à interrompre le processus d'enseignement-apprentissage. Déjà, l'évaluation sommative, de par sa nature, interrompt le processus d'enseignement-apprentissage. Quant à l'évaluation formative, on sait que certains dispositifs d'évaluation, par exemple, le mini-test, l'examen à développement, sont de nature à interrompre l'enseignement. On sait aussi qu'on peut incorporer l'évaluation formative au processus d'enseignement-apprentissage, sans l'interrompre, en utilisant des formules pédagogiques comme dispositif de cueillette d'information (Howe, 1991; Scallon, 1992). Par exemple, un professeur peut superviser du travail en équipe ou du travail individuel et faire une évaluation formative en même temps. Plusieurs autres formules pédagogiques, comme le séminaire, les réseaux de concepts, ou le journal de bord constituent d'excellentes approches d'évaluation formative sans interrompre le processus d'enseignement-apprentissage.

Selon les réponses obtenues à cette question 5, certains professeurs consacrent des proportions importantes du temps de classe à des activités d'évaluation. On peut faire l'hypothèse que, parmi leurs activités d'évaluation formative, certaines interrompent l'enseignement alors que d'autres y contribuent. De même, puisque 81% des répondants affirment que les travaux et les examens soumis à leurs étudiants contribuent "toujours" au calcul de la note inscrite au bulletin, on constate que ce qu'on appelle "évaluation formative" devient souvent, en fait, des activités d'évaluation sommative lorsqu'elles donnent lieu à des notes qui

sont versées au bulletin final. Dans la mesure où les professeurs associent l'évaluation formative à un processus qui interrompt le processus d'enseignement-apprentissage, on peut comprendre que cette conception génère des réticences à l'égard de l'évaluation formative.

Il nous apparaît souhaitable qu'une recherche future fasse l'analyse des pratiques en évaluation formative afin de distinguer les pratiques qui interrompent le processus d'enseignement-apprentissage de celles qui ne le font pas. On devrait analyser les pratiques d'évaluation formative qui se confondent à l'évaluation sommative afin de décrire les motivations et croyances des étudiants autant que des professeurs à cet égard.

L'utilisation de l'évaluation pour aider l'apprentissage

On a vu ci-haut que, dans beaucoup de cas, on évalue presque toujours de façon sommative. Et si on est conscient que l'évaluation peut aider à apprendre, on évalue quand même surtout pour gérer la classe, c'est-à-dire pour faire étudier, faire travailler et encourager à l'effort.

L'évaluation, pour un professeur sur cinq, sert peu à guider l'enseignement, surtout en sciences physiques. Selon un professeur sur dix, l'évaluation est même inutile.

On a vu que 51% des répondants croient que les résultats d'une évaluation formative ne doivent pas servir au calcul de la note inscrite au bulletin (Q49g). Dans la pratique, 56% des professeurs n'incluent que rarement (14%) ou jamais (42%), dans la note finale, des notes obtenues dans des évaluations formatives (Q50) . Il y a forte corrélation négative entre ces deux questions, l'une étant formulée négativement (τ = -,625; p = ,0001), indiquant qu'il y a une relation cohérente entre la croyance et la pratique sur ce point. Cette cohérence se raffermit lorsqu'on isole la strate Arts (τ = -,696; p = ,0001), cette strate étant celle où il y a la plus forte proportion de professeurs qui versent "toujours" à la note finale des notes d'évaluation formative. Cette situation dans la strate Arts (préuniversitaire et technique) est probablement fort normale car les

professeurs de ce groupe utilisent beaucoup, comme dispositif d'évaluation, des productions étudiantes qui, de processus à produit, servent d'abord à l'évaluation formative et ensuite à l'évaluation sommative.

Deux fonctions, une instrumentation

Il arrive que l'évaluation formative et l'évaluation sommative se confondent sur l'instrumentation mais restent distinctes dans le temps.

En effet, certaines évaluations s'appuient sur des productions étudiantes qui sont longues à compléter. Par exemple, les productions artistiques ou techniques et les travaux longs d'écriture (recherches, compositions, etc.) Dans ces circonstances, le professeur fait une évaluation formative continue à partir du projet de production de l'étudiant. Pendant le temps que dure le processus de cette production, il y a évaluation formative. À la fin de cette production, il y a évaluation sommative sur le produit de cette production.

Ainsi, sur un même dispositif d'évaluation, l'étudiant reçoit une évaluation formative pendant le processus et une évaluation sommative à la fin du processus, lorsque sa performance devient un produit. Cette évolution de la fonction évaluative dans le cadre d'une même activité correspond tout à fait à la conception qu'on doit se faire de l'évaluation formative et de l'évaluation sommative.

L'évaluation formative constitue une zone franche dans laquelle l'étudiant peut s'exercer sans crainte que ses erreurs ne lui portent préjudice. L'évaluation est formative parce que l'étudiant y reçoit, hors de portée du stress de la note, des commentaires qui l'aident à apprendre. L'évaluation sommative joue alors pleinement son rôle qui est de mesurer en quoi les buts sont atteints.

On retrouve cette situation là où l'étudiant apprend à faire quelque chose. Pendant qu'il s'exerce à faire cette chose, le professeur l'aide à corriger ses erreurs et à "faire mieux la prochaine fois". C'est l'évaluation formative.

Après un certain temps, lorsqu'en principe l'étudiant a atteint le degré de qualité voulu (qu'on appelle critères ou standards de performance), le professeur vérifie si l'étudiant sait faire cette chose et détermine à quel degré de qualité il sait le faire. C'est l'évaluation sommative.

Donnons deux exemples. En techniques de diététique, les étudiants apprennent à produire un repas de collectivité en faisant des exercices guidés. Pendant ce processus, il y a rétroaction, dialogue, régulation et remédiation. C'est l'évaluation formative. À la fin de ce processus, un repas de collectivité est produit et l'évaluation sommative s'exerce. En technique de pilotage d'avion, l'étudiant s'exerce souvent, sous la supervision de l'instructeur, à faire une approche d'atterrissage. Dans ce milieu, l'évaluation formative est nommée "debriefing". C'est là que l'étudiant et l'instructeur discutent de ce qui s'est passé "en l'air". Ils reviennent sur les techniques appropriées et sur les aspects théoriques s'il y a lieu. Plus tard, lors de l'évaluation finale par l'inspecteur du ministère des Transports, l'étudiant fera les mêmes gestes avec, en tête, les mêmes critères de performance. Cette fois, cependant, ce n'est plus une pratique (évaluation formative), c'est l'examen final (évaluation sommative).

La forte polarisation des réponses à l'énoncé 6k "L'évaluation formative, ce sont de petites évaluations qui servent à préparer à l'évaluation sommative" pourrait s'expliquer par un manque d'univocité de libellé. Certains auront répondu à partir de l'expression "petites évaluations", alors que d'autres auront choisi "préparer à". Tel que rédigé, cet énoncé aura eu au moins l'avantage de susciter des débats révélateurs entre les professeurs de deux de nos groupes réacteurs.

En discutant les réponses à cet énoncé, plusieurs professeurs ont illustré leur compréhension du concept d'évaluation formative en donnant des exemples de leurs pratiques en salle de classe. Par exemple:

Dans cet énoncé 6k, l'expression "servent à" peut être vue doublement. Les évaluations formatives peuvent servir à l'étudiant. Mais elles peuvent aussi servir à l'enseignant en lui permettant d'ajuster son évaluation sommative en fonction de la clientèle et en fonction de leurs acquis. En ce sens, l'évaluation formative peut jouer un rôle diagnostique et régulateur.

Une des constatations majeures de cette recherche est que le concept d'évaluation formative est ambigu chez plusieurs enseignants, probablement la majorité. C'est peut-être le concept le plus méconnu chez les enseignants des cégeps. Une minorité croit même que l'évaluation formative n'est pas nécessaire.

Si les spécialistes en évaluation montrent que l'évaluation peut avoir diverses fonctions, on constate qu'au collégial, elle sert avant tout à gérer la classe, à faire travailler, à faire étudier, plutôt qu'à aider à l'apprentissage. L'évaluation formative véritable, celle qui ne "compte pas au bulletin" est quasi-inexistante. Les professeurs de cégeps sont conscients de l'influence certaine de l'évaluation sur les pratiques d'étude, mais ils semblent avoir choisi de s'en servir comme levier motivateur. Plus encore, la pratique des professeurs indique que c'est surtout la note (donc un produit de l'évaluation) qui contribue à faire étudier, à motiver l'effort, à susciter la qualité des travaux, à assurer la présence en classe, et ce, pour autant qu'ils choisissent d'attribuer une note à ces "vertus".

Pour beaucoup d'enseignants, la majorité, semble-t-il, l'évaluation ne sert pas d'abord à aider le processus d'enseignement et d'apprentissage, ni à guider des décisions administratives de sanction. L'évaluation est plutôt détournée de ces fonctions et on s'en sert pour motiver les étudiants, pour suppléer, pour prévenir. Plusieurs tiennent à ce que toutes les évaluations soient sommatives à cause du pouvoir que cela leur donne sur les individus et sur le groupe et ce, même si on souhaite que l'évaluation soit simultanément "formatrice". Dans un groupe réacteur, on a entendu:

"Je fais de l'évaluation sommative mais je m'arrange pour que ça soit motivant. Je leur soumet des questions "au choix" parce que je veux leur donner une chance."

> "Même si je ne fais que de l'évaluation sommative, ça va les aider à progresser quand même parce que je donne une rétroaction corrective en classe. Mais ça compte au bulletin."

Pourtant, si l'évaluation formative et l'évaluation sommative peuvent se confondre sur l'instrumentation, on sait qu'elles ne devraient pas se confondre dans le temps. Des recherches ont démontré que, lorsqu'on tente de faire de l'évaluation formative (avec rétroaction corrective) en même temps qu'une évaluation sommative (avec note versée au bulletin final), on place les étudiants en conflit d'intérêt (Roy, 1991, p. 119). La préoccupation de l'échec ou du succès l'emporte sur la disponibilité à recevoir et à utiliser une rétroaction corrective.

Les professeurs de cégeps agissent comme s'ils n'avaient pas confiance que les étudiants, sans note, étudieraient, seraient présents activement en classe, participeraient aux exercices guidés et aux exercices autonomes. En supposant qu'ils auraient des motifs fondés de ne pas avoir cette confiance, les professeurs de cégeps agissent comme s'ils voulaient obliger les étudiants à faire ce qu'il faut pour apprendre. Or, puisque l'ordre collégial appartient au "post-obligatoire" et que les étudiants sont, à cet âge-là, libres de s'y inscrire, on est obligé de se demander comment les professeurs des cégeps conçoivent leur rôle et leurs responsabilités dans le processus d'enseignement et d'apprentissage. Pourtant, lorsqu'on propose ces arguments, certains professeurs nous répondent que, s'ils cessaient de se servir des notes pour forcer les étudiants à étudier, les taux d'échecs seraient plus élevés. Implicitement, ils laissent entendre qu'il faut contrôler les taux d'échecs et cela illustre la conception que ces professeurs se font de leur rôle ainsi que la perception qu'ils ont des attentes qu'ont, à leur égard, les administrations des collèges ou leurs collègues de département.

Analyse factorielle

Une analyse factorielle est effectuée sur l'ensemble des variables de ce champ de compétence. Cinq facteurs sont identifiés.

Facteur 1: L'évaluation peut guider l'étude. Le premier facteur regroupe trois énoncés de croyances (6b, g et h) qui, ensemble, témoignent d'une conscience que l'évaluation des apprentissages peut favoriser et guider l'étude.

Facteur 2: L'apprentissage assisté par l'évaluation. Le second facteur caractérise l'esprit de l'évaluation formative. On y trouve trois énoncés de croyances qui associent l'évaluation formative au concept de l'apprentissage assisté par l'évaluation.

Facteur 3: "Moins j'y crois, moins j'évalue". Dans ce facteur, deux énoncés de croyances (6f et 6i) sont associés à la question 5 portant sur le temps de classe consacré aux activités d'évaluation de toute nature. Ce facteur, manifestement, caractérise les professeurs qui voient l'évaluation comme une exigence administrative dont ils se passeraient bien et qui, en conséquence, pratiquent peu l'évaluation des apprentissages.

Facteur 4: Les travaux et examens vont au bulletin. Ce quatrième facteur décrit les professeurs qui ne croient pas en l'utilité de l'évaluation formative (Q6j) et qui, en toute cohérence, versent au bulletin les résultats de tous les travaux et examens (Q4).

Facteur 5: L'évaluation pour faire travailler. Dans ce dernier facteur, on retrouve deux énoncés de croyances 6a et 6k qui, associés à la question 5 portant sur le temps de classe consacré à toutes les formes d'évaluation, caractérisent les professeurs qui morcellent l'évaluation et en font beaucoup et souvent pour faire travailler les étudiants et pour les préparer à l'évaluation sommative.

Recommandations

Recommandation 1: Nous recommandons que des activités de perfectionnement soient offertes aux enseignants sur les fonctions de l'évaluation des apprentissages et, tout particulièrement, sur l'évaluation formative.

Recommandation 2: Nous recommandons que les professeurs de cégeps apprennent à mieux distinguer les diverses variables pédagogiques qui entrent en jeu lorsque l'on utilise l'évaluation, en salle de classe, pour guider la prise de décision, pour aider à l'apprentissage ou pour gérer la classe et les individus.

6.2.2 Le second champ de compétence: Les cibles de l'évaluation

L'étude des données regroupées sous ce second champ de compétence nous révèle des zones de croyances et de pratiques qui devront certes alimenter des activités de perfectionnement chez les professeurs.

Analyse

Au chapitre des croyances, on note des opinions fort partagées quant à la pertinence, au collégial, d'évaluer "autre chose que les connaissances", de même qu'à la possibilité d'évaluer les attitudes des étudiants. De plus, la question de la mesure des habiletés intellectuelles supérieures fait l'objet d'opinions très polarisées.

Mais c'est au chapitre des pratiques que nos données retiennent le plus l'attention. Dans le questionnaire, notre troisième question mesure la pratique d'attribuer des points pour certaines variables que nous considérons comme ne faisant généralement pas partie des résultats attendus de l'apprentissage. Ces variables participent à des stratégies de contrôle des étudiants. Elles servent habituellement à "encourager" l'étude ou tout autre comportement souhaité par les professeurs. Ce sont, quant à nous, des cibles de mesure qui constituent ce que plusieurs appellent un détournement de la fonction d'évaluation. Les statistiques présentées au chapitre 5 montrent l'étendue de ces pratiques.

Dans les groupes réacteurs, les statistiques observées sur cette question 3 soulèvent des commentaires souvent passionnés. Plusieurs professeurs se reconnaissent dans nos tableaux statistiques et invoquent diverses raisons pour justifier leur pratique d'incorporation de certaines de ces variables.

- certains disent que leur discipline le justifie:
  - le fait d'accorder des points pour la créativité et l'originalité se justifierait notamment, en français obligatoire et en arts. Un enseignant en Arts affirme que toutes les caractéristiques de la question 3 sont utilisées comme critères d'évaluation dans sa discipline;
  - en éducation physique, il serait justifié d'accorder des points pour les progrès réalisés en cours de sesssion;
  - en biologie, il serait nécessaire d'accorder des points pour la présence au laboratoire.
- certains ajoutent que la clientèle étudiante le justifie, étant trop peu motivée de façon intrinsèque.

En situation de salle de classe, des enseignants affirment que la présence en classe a moins de valeur que la qualité de cette présence et qu'ils auront tendance à tenir compte de la qualité de la présence des étudiants, du degré de participation en classe, pour éventuellement accorder une note de passage.

Dans des laboratoires, par exemple en chimie, des professeurs disent attribuer 10% de la note à la qualité de la participation au laboratoire: participation active, effort, qualité du travail en équipe.

Notre énoncé 3c), "l'exécution de devoirs ou d'exercices à la maison ou en classe" en a surpris quelques-uns. En effet, certains professeurs se disent déconcertés de nous voir classer cette pratique dans le questionnaire, comme un "à côté", au même titre que "la présence assidue aux cours". Ces professeurs disent considérer qu'il va de soi qu'il faut donner des points pour l'exécution de devoirs ou d'exercices à la maison ou en classe. Leur opinion est soutenue par d'autres mais ne fait pas l'unanimité. Quant à nous, ceci soulève la question de la fonction pédagogique des devoirs et des exercices, selon qu'ils constituent des exercices contribuant au processus de l'apprentissage ou des dispositifs

d'évaluation de la compréhension. À cela, plusieurs professeurs ajoutent qu'ils croient fermement que les étudiants ne feront pas les devoirs et les exercices s'ils ne "comptent pas".

Analyse factorielle

L'analyse factorielle des variables appartenant au second champ de compétence a révélé sept facteurs composés d'au moins deux variables.

Facteur 1: Le contrôle des comportements. Ce premier facteur caractérise spécifiquement les professeurs qui ont affaire à mesurer l'atteinte d'objectifs de savoir-faire (Q2c) et de savoir être (Q2g). Selon ce facteur, ce sont ces professeurs qui, du même coup, ont le plus tendance à accorder des points pour la motivation et l'effort (Q3a), la présence assidue aux cours (Q3b), la participation en classe (Q3e) et les progrès réalisés au cours de la session.

Facteur 2: Les habiletés intellectuelles. Le second facteur associe la mesure des habiletés intellectuelles supérieures (Q2d, e, f) à l'attribution de points pour la créativité et l'originalité.

Facteur 3: Valoriser les connaissances. Ce troisième facteur est composé de croyances. Il caractérise les professeurs qui valorisent d'abord la mesure de la connaissance de la matière.

Facteur 4: Mesurer les habiletés de base. Ce facteur décrit la pratique d'évaluer uniquement la connaissance et la compréhension.

Facteur 5: Tout évaluer. Ce facteur est celui qui caractérise la croyance à l'effet qu'il faut évaluer tout ce qui est enseigné.

Facteur 6: Valoriser les habiletés supérieures. Ce sixième facteur est, pour ainsi dire, à l'opposé du facteur trois. Il décrit ceux qui croient qu'il faut évaluer des habiletés supérieures à la connaissance.

Facteur 7: Selon ce dernier facteur, plus le professeur est seul à choisir les cibles de ses évaluations, plus il a tendance à donner des points pour le respect de ses exigences de présentation de travaux et pour l'exécution de devoirs à la maison.

Recommandation

Recommandation 3: Les professeurs devraient savoir reconnaître les divers types d'apprentissages qui sont présents dans les objectifs qu'ils définissent et devraient savoir mesurer ces apprentissages. À cette fin, des activités de perfectionnement devraient être offertes aux enseignants sur les dispositifs les plus appropriés à la mesure des divers types d'apprentissages attendus des étudiants.

Les notes qui sont publiées sur les bulletins des cégépiens doivent être valides et interprétables. Étudiants, parents et employeurs doivent pouvoir lire des notes univoques et leur accorder toute la crédibilité sur laquelle se fonde le processus de sanction des études. C'est pourquoi les professeurs ne doivent pas polluer la note de leurs étudiants en y incorporant des variables qui n'ont rien à voir avec les résultats de l'apprentissage. Les professeurs doivent éviter de faire du détournement de l'évaluation sommative en utilisant celle-ci comme moyen de gérer la classe ou comme levier pour faire étudier et pour faire travailler.

Recommandations

Recommandation 4: Les professeurs doivent savoir reconnaître les pratiques d'évaluation qui constituent du détournement de l'évaluation et qui donnent lieu à des notes polluées. À cette fin, les professeurs doivent savoir résoudre leurs problèmes de gestion de classe ou de contrôle des comportements des étudiants sans compromettre l'univocité et l'interprétabilité des notes transmises au bulletin.

Les administrateurs de collèges doivent donner aux professeurs des moyens concrets pour résoudre ces problèmes en proposant notamment, des règles de conduite et en partageant avec eux la responsabilité de la gestion ou des comportements en classe et dans l'établissement.

Recommandation 5: Les politiques institutionnelles d'évaluation des apprentissages ainsi que la formation offerte aux enseignants devraient inciter ces derniers à faire en sorte que les notes finales soient cohérentes avec les objectifs d'apprentissage et ne soient pas polluées de diverses variables non pertinentes aux apprentissages prévus.

6.2.3 Le troisième champ de compétence: Les qualités de l'évaluation

Nous savons qu'une évaluation de qualité correspond aux attributs suivants:

- elle porte sur une cible explicite, spécifique, bien comprise et à laquelle correspond une méthodologie, un dispositif approprié;
- les sources d'erreur de mesure sont contrôlées. Certaines caractéristiques des étudiants, le processus de mesure, l'environnement de mesure sont des exemples de sources d'erreur de mesure. Elles sont externes aux compétences mesurées mais elles peuvent influencer les résultats;
- l'échantillon mesuré est représentatif des compétences ciblées. Cette représentativité permet de faire des inférences valides et fiables;
- l'évaluation fournit une information compréhensible et utilisable, en fonction des buts recherchés.

L'analyse des données et les témoignages dans les groupes réacteurs montrent que ces attributs ne sont pas toujours présents au cégep.

Analyse

Planification de l'évaluation: Les banques de questions d'examen

Beaucoup d'enseignants peuvent réutiliser des questions d'examen d'une session à l'autre, soit en empruntant des questions à des collègues, soit en se donnant des banques de questions d'examen (Q15). Ces banques peuvent prendre diverses formes et la documentation spécialisée décrit les méthodologies appropriées en plus de démontrer qu'il y a plusieurs avantages à réutiliser des questions d'examen:

1. Cela peut diminuer considérablement le fardeau de la tâche d'évaluation;

2. Les professeurs peuvent s'inspirer d'anciennes questions pour en générer de nouvelles;

3. Ils peuvent améliorer leurs questions d'une session à l'autre, au vu des réponses obtenues lorsqu'ils les administrent;

4. Ils peuvent se prévaloir d'un processus d'analyses d'items pour corriger et raffiner les consignes, énoncés, choix de réponses, barêmes de correction et ce, à partir de divers critères d'analyses, dont: la difficulté, la capacité de discrimination, la plausibilité;

5. Ils peuvent de même corriger et enrichir les consignes de questions à développement long (questions "ouvertes") ou des travaux longs (recherches, projets, protocoles de laboratoires, dissertations, etc.)

Ces processus peuvent grandement contribuer à améliorer la validité de leurs évaluations. Pourtant, 22% des enseignants ne réutilisent que très rarement, sinon jamais, des questions déjà existantes et près de 40% n'y ont recours que parfois.

Plusieurs raisons peuvent amener des professeurs à ne pas préparer la plupart de leurs questions d'examen à partir de questions déjà existantes (Q15). Notamment, certains peuvent craindre le plagiat, et le phénomène par lequel les étudiants se transmettent des questionnaires d'une session à l'autre. Mais, s'il est bien connu que des étudiants peuvent chercher divers moyens pour faire bonne figure aux examens, il est loin d'être certain qu'ils se serviraient systématiquement des questions d'examen d'une session antérieure pour guider leur étude.

D'autres raisons peuvent expliquer que des professeurs ne préparent pas la plupart de leurs questions d'examen à partir de questions déjà existantes (Q15). Certains n'ont que très rarement, sinon jamais, le même cours à enseigner d'une session à l'autre. Il leur est alors inutile de se faire des banques d'items. D'autres, enfin, ne le font pas par manque

d'organisation ou par méconnaissance des avantages qu'ils retireraient à le faire.

Une planification contrôlée

Nos données quantitatives montrent que certains professeurs éprouvent des problèmes de planification et d'organisation de l'évaluation (Q8 et 15).

Un enseignant qui a bientôt 10 ans d'expérience en salle de classe nous dit qu'il attend la veille de l'examen pour écrire ses questions d'examens, parce qu'il ne sait pas d'avance où il sera rendu dans son contenu à la date de l'examen. Ça dépend de ses étudiants. Ce professeur nous apprend qu'il peut consacrer plus ou moins de temps à certains chapitres de son cours, selon les difficultés qu'éprouvent les étudiants à comprendre. Mais puisque le calendrier scolaire n'est pas souple, ce sera la planification de ce professeur qui devra être souple. Cela implique qu'il est possible que certains contenus ne seront pas "vus" ou ne le seront que partiellement. Certains objectifs ne seront peut-être pas atteints. Ce professeur éprouve manifestement des difficultés à garder du contrôle sur la planification de son cours. Peut-être a-t-il mal diagnostiqué les compétences préalables de ses étudiants en fonction des objectifs qu'il vise dans son cours ou en fonction des approches pédagogiques qu'il utilise. Ce témoignage illustre bien que des pratiques incorrectes peuvent résulter d'un manque de compétence à planifier le déroulement du cours, à contrôler la cadence de ce déroulement, à choisir les moyens de remédiation les plus efficaces dans les circonstances, à planifier les stratégies et les dispositifs d'évaluation.

Contrôle de la validité

Sur le thème du contrôle de la validité, les questions 16 et 17 du questionnaire nous apprennent qu'une majorité des répondants ne relisent pas toujours le contenu ou les objectifs de leur cours lorsqu'ils préparent

leurs questions d'examen. Parmi les réactions reçues d'un groupe réacteur, un professeur s'inquiète que nous ne tirions des conclusions hâtives de nos statistiques sur le contrôle de la validité des évaluations. Selon lui, il arrive que des professeurs enseignent un même cours pendant plusieurs sessions consécutives et en arrivent à le maîtriser si bien qu'il n'est plus nécessaire de relire le contenu ou les objectifs pour préparer les examens. À cette remarque, un autre professeur dans la salle met en doute ce raisonnement en disant que si un professeur en est rendu à ne plus avoir besoin de relire son plan de cours, c'est peut-être signe qu'il serait urgent qu'il le fasse.

Préparation des étudiants

Le quart des répondants ne préviennent pas toujours leurs étudiants à l'avance de ce sur quoi porteront leurs examens (Q18) et 60% ne donnent pas toujours les critères de correction à l'avance (Q9). Les témoignages de certains enseignants, en groupe réacteur, donnent des éclairages susceptibles d'expliquer ces pratiques. Ils laissent entendre que ces pratiques procèdent peut-être, au moins dans certains cas, d'un calcul stratégique orienté vers le contrôle de l'étude. Ils affirment que leurs étudiants n'étudieront pas l'ensemble de la matière s'ils savent que l'examen portera sur un aspect plutôt qu'un autre. Ces professeurs semblent croire qu'il vaut mieux laisser les étudiants dans le doute quant au contenu de l'examen, les obligeant à étudier et à réviser l'ensemble du contenu. Cette croyance n'est pas mesurée dans notre recherche mais devrait, à notre avis, être investiguée, à la lumière, d'une part, des recommandations d'experts sur le principe du secret quant au domaine mesuré en examen, de même que, d'autre part, des recherches sur les comportements d'étude des étudiants des cégeps.

Administration de l'évaluation

Nous avons vérifié si les répondants qui écrivent souvent ou parfois les questions d'examen au tableau ou les projètent sur acétates (Q20a) les font aussi imprimer pour en distribuer un exemplaire à chacun des étudiants. Un croisement des réponses montre que ce n'est pas le cas. Ceux qui écrivent les questions d'examen au tableau ou les présentent en acétates ne sont généralement pas les mêmes répondants qui disent distribuer un questionnaire imprimé.

Les données recueillies montrent que cette pratique est heureusement très minoritaire. Elle concerne 5% de l'ensemble des répondants (2% souvent et 3% parfois; personne n'a répondu "toujours"). Ces chiffres peuvent sembler rassurants à quiconque se préoccupe des conséquences de cette pratique sur l'erreur de mesure. Mais, si on se souvient que le réseau collégial compte 12 000 professeurs, on calculera que 5% de la population enseignante correspond à 600 professeurs, soit, en moyenne, une douzaine d'enseignants par collège.

La pratique d'écrire les questions au tableau ou de les projeter sur acétates (Q20a) pourrait être attribuable à des lacunes de planification, menant certains professeurs à préparer leur examen à un moment tellement rapproché de l'examen lui-même qu'ils n'ont plus le temps d'en faire un texte écrit. D'autres professeurs peuvent choisir cette pratique parce qu'elle permet de la spontanéité dans l'évaluation. En effet, un professeur participant à un groupe réacteur a chaudement défendu sa pratique d'écrire ses questions d'examen au tableau. Dans la dynamique des échanges avec ses étudiants, un thème de discussion peut émerger et ce professeur affirme qu'il lui convient de décider que ce thème sera traité par écrit, sur le champ, par ses étudiants et que cela constituera une évaluation sommative. Selon cet enseignant, proscrire la rédaction au tableau de questions d'examen, ce serait limiter la spontanéité et la créativité en évaluation, surtout dans les disciplines dont le contenu n'est pas obligé.

Un autre enseignant affirme qu'il écrit toujours ses questions au tableau parce qu'il craint le plagiat: il ne veut pas que ses questions d'examen circulent entre les groupes et entre les sessions car il veut pouvoir s'en servir plusieurs fois. Et, selon lui, il n'y a de toute façon aucune différence entre imprimer les questions et les écrire au tableau.

D'autre part, chez la majorité de professeurs qui font toujours imprimer et distribuer un questionnaire d'examen (Q20b), une faible proportion, 4%, fait aussi la lecture à haute voix. D'ailleurs, la grand majorité (89%) de tous ceux qui font la lecture à haute voix de leurs questionnaires d'examens impriment aussi ces questionnaires.

Il peut étonner que des professeurs présentent doublement leurs questionnaires d'examen, par écrit et oralement. Peut-être est-ce là l'expression de la volonté de certains d'être les plus aidants possible. Peut-être y a-t-il là quelques enseignants qui doutent, à tort ou à raison, de la clarté de leurs questions écrites. Cette hypothèse a trouvé preneur dans un groupe réacteur. Un participant nous dit qu'il imprime ses questionnaires mais les lit aussi en classe. Il les paraphrase car il pense que certains énoncés peuvent manquer de clarté. Il lit ses questions à haute voix afin de les interpréter, au cas où certains étudiants ne les comprendraient pas. De toute façon, dit-il, sa "calligraphie est abominable". Mais la lecture à haute voix des questions lui sert aussi à les commenter car il veut dire aux étudiants comment ils doivent répondre à la question. Il en profite donc pour leur donner des conseils méthodologiques pour résoudre les problèmes posés, pour les guider dans leur réponse.

Cette pratique, même si elle est minoritaire, mérite qu'on s'y attarde dans un processus de perfectionnement à cause des biais que peut causer une interprétation orale d'un questionnaire d'examen imprimé.

Les questions "au choix"

Les tableaux du chapitre 5 montrent qu'un professeur sur quatre soumet parfois ou souvent des questions "au choix" (Q24). Dans un groupe réacteur, un professeur décrit sa pratique à cet égard. Il utilise des questions à développement à "livre ouvert", afin de faire travailler les étudiants. Et il laisse répondre à des questions au choix afin de motiver les étudiants. Ce professeur écrit, disons, six questions et demande à ses étudiants de répondre, au choix, à deux de ces questions. Selon lui, cela leur permet de travailler sur des thèmes qui les intéressent. Cela leur donne une chance, aussi: c'est peut-être leur dernière chance de comprendre car, après l'examen, ce professeur passe à la suite de son contenu.

Ainsi, certains professeurs semblent se servir des questions à développement, particulièrement des questions "au choix", comme stratégie de motivation et, surtout, comme exercice guidé, favorisant l'apprentissage des concepts.

La validation a posteriori

On sait que 94% des répondants croient qu'après la passation d'un examen, il est utile de faire l'analyse des questions pour en évaluer la qualité (Q33a). Cependant, en croisant leurs réponses avec la question 14 (τ = ,402 ; p = ,0001), on remarque que ceux-là ne font pas toujours l'analyse de leurs questions d'examen, après la correction. De ceux qui croient qu'il est utile de faire l'analyse des items d'examen, 43% le font rarement. De même, chez ceux qui se disaient très en accord avec ce principe, 13% ne s'y adonnent que parfois ou rarement et 38% disent le faire souvent.

De plus, chez ceux qui croient qu'il est utile de faire l'analyse des questions a posteriori, on constate aussi qu'ils ne demandent pas tous l'opinion des étudiants sur la formulation des items du questionnaire (Q22). 40% seulement de ceux qui sont très en accord et 23% de ceux qui sont en accord avec le principe de l'analyse des questions a posteriori

vont toujours ou souvent demander à leurs étudiants leur opinion rétrospective sur la formulation des items du questionnaire.

D'aucuns diront que la tâche est lourde et que, malgré la conscience de l'importance de faire l'analyse des questions d'examen, on ne peut pas toujours la faire. On manque de temps ou on n'a pas les moyens techniques de la faire. À cet égard, la question 58-6 était destinée à vérifier si les professeurs estiment qu'il leur serait utile qu'on mette à leur disposition un système informatisé d'analyse d'items. Seulement 26% des répondants ont identifié cette ressource comme éventuellement utile, dont moins du tiers de ceux qui croient qu'il est utile de faire des analyses d'items (Q33a) et qui en font souvent ou toujours (Q14). Ainsi, la validation a posteriori des questions d'examen, donc après la correction des examens, n'est pas une pratique de la majorité des professeurs. Plusieurs n'en font pas ou peu, même s'ils croient que c'est important. Et ils ne souhaitent pas beaucoup qu'on leur offre des ressources qui les aideraient à ce chapitre. Diverses hypothèses pourraient être explorées pour essayer d'expliquer que ces pratiques ne soient pas généralisées: la lourdeur de la tâche, la congestion du curriculum, etc. Il faudrait aussi explorer l'hypothèse du manque d'information sur ce processus. Il nous apparaît probable que plusieurs enseignants ne connaissent pas les principes et les mécanismes de l'analyse des items et, de là, ne voient pas comment cela peut leur être utile.

L'écart évident entre les pratiques d'analyse d'items (Q14) et les demandes de ressources à cet égard (Q58-6) peut s'expliquer, selon une enseignante d'un groupe réacteur, par le fait que les ressources informatisées en analyse des items ne peuvent s'appliquer qu'aux questions à choix multiples, alors que beaucoup d'enseignants utilisent des questions à développement. Ces derniers font vraisemblablement l'analyse des items par un procédé subjectif, fondé sur les profils de réponses de leurs étudiants. Un autre ajoute qu'il est capable d'estimer la qualité de ses questions à partir des réponses des étudiants et de sa connaissance générale des capacités de ses groupes d'étudiants. Il peut, sans aide, interpréter les résultats des examens.

Plusieurs professeurs font en classe un retour sur les résultats de leurs examens et échangent des commentaires avec leurs étudiants sur le contenu mesuré dans l'examen, de même que sur la compréhension des questions. Un effet de cette pratique, c'est que l'on ne requestionne pas l'énoncé des questions d'examen lorsque les réponses des étudiants se sont avérées bonnes en général. Spontanément, ces questions d'examen sont alors considérées comme correctement écrites: "celle-là a bien marché". Ces professeurs affirment que, d'intuition, ils peuvent alors distinguer entre une "question qui a bien marché" parce qu'elle était trop facile et une "question qui a bien marché" parce que les étudiants ont bien appris. Il y a un autre motif qui pourrait amener des professeurs à ne pas demander d'aide pour l'analyse des items, et ce motif tiendrait à la notion d'ingérence. Selon une enseignante d'un groupe réacteur, certains professeurs pourraient craindre qu'en demandant une aide professionnelle, ils soient évalués dans leur compétence pédagogique ou, spécifiquement, dans leur compétence en évaluation. Aussi, des professeurs ne demanderont pas d'aide extérieure pour améliorer leurs examens, soit parce qu'ils pensent pouvoir se débrouiller seuls, soit parce qu'ils craignent une éventuelle ingérence.

La crédibilité des évaluations, du point de vue des professeurs

Dans l'ensemble, les professeurs accordent une crédibilité assez élevée aux dispositifs de mesure qu'ils utilisent. Selon eux, les résultats des évaluations sont plutôt fiables et représentatifs. De même, nos répondants sont généralement en accord avec les critères de validité de la mesure, proposés en énoncés de croyances.

Cette confiance générale est cependant entachée de certaines réserves que nous jugeons fort sérieuses. D'une part, près de 30% des répondants affichent des doutes sur la représentativité des résultats (Q33e).

D'autre part, deux professeurs sur cinq, sans différence entre les strates, pensent qu'il y a manque de rigueur dans les pratiques d'évaluation de

beaucoup d'enseignants (Q33h). Sur ce point, d'ailleurs, même le taux de "ne sais pas" est très élevé. En effet, 38% des répondants, soit deux autres professeurs sur cinq ne savent pas s'il y a ou non manque de rigueur. Le tableau C3-a, au chapitre 5, illustre bien l'ampleur des limites de la crédibilité que nos répondants accordent à la rigueur des pratiques d'évaluation. Si l'on fait confiance à ce que l'on fait soi-même en évaluation, on n'accorde pas nécessairement la même crédibilité aux pratiques évaluatives des autres professeurs ainsi qu'à leurs résultats. Seulement 20% des répondants croient que les enseignants ne manquent pas de rigueur en évaluation.

Ces résultats suscitent plusieurs réactions dans les groupes réacteurs. Quelques-uns ont centré leur commentaire sur la compréhension de notre énoncé de question. Certains pensent que des répondants ont pu être incertains du sens à donner à l'expression "manque de rigueur", ce qui pourrait contribuer à expliquer, au moins partiellement, le taux très élevé de "ne sais pas". Par ailleurs, selon certains professeurs, des répondants ont peut-être répondu en pensant à leurs propres pratiques et en transposant sur leurs collègues. Ainsi, certains se seraient dits en accord ou en désaccord à partir d'une évaluation de leurs propres pratiques. Mais de façon générale, les professeurs trouvent que nos répondants font une critique sévère de la qualité des pratiques d'évaluation ou de notation. Selon ces professeurs, plusieurs enseignants expriment leur désaccord à ce qu'un collègue ait "fait passer" tel ou tel étudiant. D'autres expriment un certain ressentiment à l'égard de collègues apparemment plus souples, sinon laxistes, sur la durée de la session, le type de travaux prescrits ou encore le type ou contenu des examens. Ces professeurs disent que les pratiques des autres les exposent aux pressions des étudiants qui les jugent alors trop rigoureux dans leur pédagogie ou dans la notation. Ils craignent d'être considérés comme mauvais professeurs ou comme professeurs trop sévères à cause, par exemple, d'une moyenne de fin de session inférieure à celle des collègues.

Selon les professeurs des groupes réacteurs, les répondants expriment de l'exaspération ou de la frustration parce qu'ils côtoient, sans espoir de changer les choses, des collègues qui, disent-ils, nuisent à l'image

professionnelle des enseignants et contribuent à dévaloriser la profession. Une enseignante interprète que le fort pourcentage d'accord à cet énoncé traduit peut-être l'existence de préjugés entre les strates. Des professeurs d'une strate pourraient avoir répondu à cette question en pensant aux professeurs d'une autre strate.

En somme, les participants aux groupes réacteurs pensent que cette question constitue un jugement de l'ensemble des pratiques en évaluation. Et plusieurs interprètent que, dans les circonstances, "beaucoup" est l'équivalent de "beaucoup trop".

Pourtant, certains semblent avoir acquis une tolérance aux lacunes dans la qualité de l'évaluation. Un commentaire reçu en groupe réacteur est très éloquent sur ce point:

> "Il y a des professeurs qui évaluent très mal et il y a des professeurs qui évaluent très bien. Et en fin de DEC, dans l'ensemble, les notes qu'un étudiant aura à son bulletin seront quand même représentatives de ses compétences réelles. Comme si on faisait la moyenne de bonnes et de mauvaises évaluations".

Analyse factorielle

L'analyse factorielle des variables associées à ce troisième champ de compétence dégage dix facteurs signifiants, décrits ci-après. Trois autres facteurs sont considérés comme non signifiants et non utiles, étant issus, dans deux cas, d'associations incongrues et dans un cas, d'une variable unique.

Facteur 1: Ce premier facteur est celui de la rigueur dans l'établissement des critères d'évaluation, et dans la diffusion de ces critères auprès des étudiants (Q7, 9, 17 et 33i).

Facteur 2: Le second facteur décrit l'enseignant qui planifie ses évaluations et qui en fait une validation a priori auprès de collègues (Q13 et 31).

Facteur 3: Ce facteur est caractérisé par trois croyances, toutes axées sur la valorisation des démarches de validation des questions d'examens (Q33 a, b, c).

Facteur 4: Ce facteur est celui de l'examen à développement administré dans des conditions indulgentes. C'est ici qu'on retrouve la pratique de proposer des questions "au choix", de permettre de répondre à la maison, d'avoir accès à toute documentation utile (Q24, 25 et 26). Notons qu'en principe il est tout à fait souhaitable que les étudiants aient accès à toute documentation utile dans plusieurs contextes d'évaluation, particulièrement lorsqu'on veut mesurer des habiletés intellectuelles supérieures. Dans ces cas là, nous reconnaissons qu'il ne s'agit pas d'une pratique fondée sur l'indulgence. Ici, cependant, la variable Q26 est incluse par l'analyse factorielle dans ce facteur caractérisant des pratiques d'administration d'examens à développement qui "donnent des chances" au répondant.

Facteur 5: Ce sixième facteur est celui de la banque d'items. Il caractérise la pratique de conserver et de réutiliser les questions d'examen (Q15 et 29). Cette pratique est associée à une économie de temps au chapitre de la planification et de la rédaction des questions (Q10).

Facteur 6: Ce facteur caractérise la pratique de la participation étudiante tant à la rédaction des questions (Q19b) qu'à la rétroaction sur le contenu de l'examen (Q32c) ou sur la forme de l'examen (Q22).

Facteur 7: Deux croyances caractérisent ce facteur, centré sur la préoccupation d'équité et de représentativité des évaluations. Les croyances 33d et f concernent la nécessité d'utiliser les mêmes critères pour tous les étudiants ainsi que l'importance de mesurer plusieurs fois afin d'obtenir des évaluations fidèles.

Facteur 8: Selon ce facteur, la pratique de bien préparer les étudiants en vue de l'examen (Q18 et 19a et c) est associée à la croyance que les résultats de l'évaluation sont représentatifs et signifiants (Q33g et corrélation négative avec Q33e).

Facteur 9: Ce facteur décrit une planification de l'évaluation qui prend en compte l'importance relative des objets évalués afin que la note soit représentative des objectifs maîtrisés.

Facteur 10: Les pratiques visant à rendre les items cohérents avec le contenu visé, autant lors de la rédaction des questions que lors de l'analyse a posteriori, caractérisent ce dernier facteur.

Recommandations

Recommandation 6: Nous recommandons que les professeurs développent et maintiennent des compétences qui favorisent les pratiques les plus appropriées au chapitre des qualités de l'évaluation des apprentissages. Les professeurs devraient se donner l'occasion de valider ou de mettre à jour leurs croyances et leurs pratiques touchant notamment: la planification de l'évaluation, le contrôle de la qualité, a priori, le contrôle de la validité, l'administration de l'évaluation, la préparation des étudiants, la correction et notation, les améliorations a posteriori. Cette formation devrait leur permettre de fonder des pratiques qui donnent lieu à des évaluations équitables, valides et représentatives, desquelles ils pourront tirer des inférences fiables et signifiantes.

Recommandation 7: Les enseignants devraient comprendre les biais et les erreurs de mesure que peuvent générer diverses pratiques d'administration des examens, dont la pratique de laisser les étudiants répondre à des questions au choix ainsi que la paraphrase d'un questionnaire imprimé. De plus ils devraient apprendre à rédiger des questions équivalentes en termes de difficulté, de cibles, de niveau taxonomique, ainsi que représentatives des objectifs visés.

6.2.4 Le quatrième champ de compétence: Les instruments d'évaluation

L'analyse factorielle n'a pas révélé, dans ce groupe de variables, de facteurs signifiants et utiles. En effet, dans ce champ de compétence, la plupart des variables concernant les pratiques sont tributaires des dix-sept dispositifs de mesure proposés dans le cadre des questions 34 à 41. À cause de cela, la plupart des facteurs extraits sont composés de groupes de dispositifs de mesure et nous considérons que l'analyse factorielle est ici sans valeur.

Analyse

Les questions 12 et 30 ont mesuré les proportions d'enseignants qui utilisent l'examen ou le travail écrit pour effectuer l'évaluation des apprentissages de leurs étudiants. Mais ces deux grandes familles de dispositifs de mesure ne conviennent pas à toutes les strates. En effet, on voit au tableau P4-b, en annexe 5, que les professeurs d'Arts (pré-universitaire et technique) utilisent peu le travail écrit de même que l'examen. Ces professeurs utilisent manifestement d'autres dispositifs d'évaluation. La question 34 montre bien que les profils d'utilisation des divers dispositifs de mesure varient d'une strate à l'autre.

Parmi ceux qui font usage de l'examen, la question 21 nous révèle que les professeurs utilisent d'abord la question à développement court, puis les questions à choix multiples et, en dernier, les questions à développement long. Les choix des professeurs à cet égard sont fonction de plusieurs variables: la discipline, les habiletés ciblées, les traditions disciplinaires. On a vu, au chapitre 2, que les professeurs choisissent souvent d'utiliser un type d'items plutôt qu'un autre en fonction du type d'étude qu'ils veulent "déclencher" chez leurs étudiants. Mais il arrive, semble-t-il, que ces pratiques soient fondées sur des habitudes ou des traditions disciplinaires. C'est dans ce contexte, d'ailleurs, qu'une enseignante, en entrevue, a commenté sévèrement l'attitude de ses collègues:

> "Si les professeurs de français se donnent tant de mal à corriger sans cesse pendant des heures, c'est qu'ils manquent d'imagination. Et si on leur propose de changer des choses dans leur façon de faire, ils répondent qu'on ne peut rien changer. Ils ne rationalisent pas leur conception de leur tâche."

Pour d'autres, le choix des types d'items peut être fondé sur des variables psychologiques comme la recherche d'une objectivité sécurisante. Par exemple, en entrevue de pré-consultation, une enseignante de la strate Sciences humaines nous fait l'observation suivante:

> "Les professeurs qui utilisent des questions à développement doivent apprendre à écrire des consignes claires. Sinon, ils risquent de se rendre compte, en corrigeant des piles de travaux de 4 pages, qu'ils ne sont pas comme ils le voulaient et que le barème de correction s'applique mal, parce que les consignes ne sont pas claires. Il y a une zone grise dans les questions à développement qui rend les professeurs vulnérables à des contestations de notes des étudiants. Plusieurs professeurs vivent des sentiments d'angoisse pendant la correction. Ils doutent d'eux-mêmes à cause du sentiment qu'ils sont subjectifs dans leur jugement. Ils sont mal à l'aise de donner une mauvaise note basée sur des jugements subjectifs. À cause de ça, plusieurs professeurs se réfugient dans les questions à choix multiples pour avoir plus de sentiment d'assurance dans leur verdict. Les questions à choix multiples leur donne une impression d'objectivité."

Dans cette série de variables, les questions sur les croyances portent sur des dimensions plutôt techniques. Elle impliquent que les répondants possèdent des connaissances minimales en mesure et évaluation, soit pour donner une réponse autre que "ne sais pas", soit pour afficher, comme groupe, des consensus à l'égard de ces dimensions techniques.

Or, la lecture des tableaux statistiques du chapitre 5 sur les instruments d'évaluation illustre, croyons-nous, la méconnaissance des professeurs sur plusieurs aspects en mesure et évaluation. En effet, dans toute la série d'énoncés de croyances 33k jusqu'à 33s, il y a plus de 5% de "ne sais

pas". La proportion monte à 10% lorsque vient le temps de répondre aux énoncés traitant notamment de l'évaluation formative et des capacités de mesure des questions à développement et des questions à choix multiples. Cette proportion de "ne sais pas" touche le seuil de 15% sur les concepts d'examens à livre ouvert et d'examens synthèses et atteint le 21% (un professeur sur 5) sur le concept de "choix au hasard".

Un tableau de contingence valide (τ = -,162; p = .0001) entre les questions 33k et 33r illustre la méconnaissance de nombreux répondants. On y voit que, de ceux qui ne savent pas si les questions à développement peuvent seules mesurer les apprentissages de haut niveau cognitif (33k), 48% ne savent pas non plus si les questions à choix multiples peuvent faire de même (33r).

L'étude des répondants qui "ne savent pas" a révélé une différence significative ($p < .05$) chez les professeurs qui utilisent ou n'utilisent pas l'exposé magistral formel comme formule pédagogique (Q66-2). En effet, chez les professeurs qui enseignent par exposé magistral formel, les opinions sont plus nettes sur la question de savoir si les questions à développement peuvent, seules, mesurer les apprentissages de haut niveau. Chez ces professeurs, la proportion de "ne sais pas" est de 6%, alors que chez les professeurs qui n'utilisent pas l'exposé formel, la proportion de "ne sais pas" est de 11%. Il y a certes lieu de se demander - et ce pourrait être un sujet de recherche future - quelle est la relation entre cette formule pédagogique et le fait d'avoir une opinion moins incertaine sur la question de la mesure des apprentissages de haut niveau.

Les examens "à livre ouvert"

Le concept d'examens "à livre ouvert" fait l'objet de croyances peu favorables à une pratique fréquente. Pourtant, un professeur appartenant à la strate Administration affirme que, dans sa discipline, plusieurs évaluations se font "à livre ouvert" car on y mesure non pas les connaissances mémorisées, mais plutôt la capacité de résoudre des problèmes. Ce professeur ajoute même que l'une des habiletés mesurées consiste à pouvoir

se servir de la documentation publiée sur les logiciels afin d'y trouver des éléments de réponse.

L'autoévaluation

Plusieurs auteurs font l'éloge des avantages pédagogiques et des vertus éducatives de l'autoévaluation: Campbell et Felx (1988); Doyon et Juneau (1991); Nunziati (1991); Tousignant et Morissette (1990). Dans la perspective de la planification de la formation en évaluation, des chercheurs comme Gong et Venezky (1992) affirment que l'on devrait encourager les professeurs à permettre à leurs étudiants de faire l'autoévaluation de leurs apprentissages.

Pourtant, les croyances de nos répondants favorisent très peu la participation de leurs étudiants à l'autoévaluation de leurs apprentissages (Q33m et 33n). De même, l'autoévaluation est utilisée par 12% seulement des professeurs (Q34) dont, tout particulièrement, les professeurs de techniques biologiques (voir à cet effet le tableau P4-b en annexe 5). Paradoxalement, l'autoévaluation se situe au troisième rang des dispositifs d'évaluation que les professeurs utiliseraient davantage s'ils le connaissaient mieux (Q37).

On a vu que les professeurs des cégeps utilisent une grande variété de dispositifs et d'approches en évaluation. Plusieurs de ces dispositifs leur sont mal connus et le lecteur trouvera très révélateur de revoir les réponses que nous avons obtenues aux questions 34 à 41, notamment à la question 38 "Laquelle de ces approches utiliseriez-vous davantage si vous aviez plus de temps?" et à la question 37 "Laquelle de ces approches utiliseriez-vous davantage si vous la connaissiez mieux?" Sur ce dernier point, on notera avec intérêt que les professeurs souhaitent mieux connaître, en toute priorité, les techniques suivantes: le schéma ou réseau de concepts, la résolution de problèmes, l'autoévaluation, le dossier anecdotique et le journal de bord, l'observation avec grille d'appréciation ou check-list.

Tenant compte aussi des malaises qui ont été exprimés au sujet de la planification de l'évaluation ainsi que des critères d'évaluation et de notation, nous pensons que les efforts de formation des professeurs devraient tenir compte du besoin de pouvoir utiliser une variété de dispositifs de mesure.

Recommandations

Recommandation 8: Les professeurs devraient savoir utiliser une instrumentation variée afin de pouvoir évaluer, notamment, des habiletés et connaissances procédurales ainsi que des attitudes. La formation devrait cibler, en priorité, les dispositifs de mesure les plus souvent invoqués lors des questions 37 et 38, à savoir ceux qui seraient utilisés davantage s'ils étaient mieux connus ou si les professeurs avaient plus de temps.

Les pratiques sont reliées aux disciplines enseignées et les départements se sont donnés, à cet égard, des cultures propres. Les cultures internes des disciplines ou des départements favorisent l'adoption par les professeurs de certains des dispositifs de mesure proposés à la question 34. De même, d'autres dispositifs sont exclus. Ces pratiques, souvent fondées sur des croyances propres aux disciplines, ont des conséquences sur la tâche des professeurs de même que sur les choix didactiques, en salle de classe. Dans certains cas, les choix de dispositifs de mesure peuvent avoir des conséquences défavorables sur la validité de la mesure.

Recommandation 9: C'est pourquoi nous recommandons que les professeurs soient invités à explorer la pertinence de diversifier les dispositifs d'évaluation qu'ils utilisent et ce, au-delà des cultures disciplinaires. Pour ce faire, les croyances à l'égard des dispositifs d'évaluation devraient être identifiées clairement et les fausses croyances devraient être discutées de façon à permettre aux professeurs de faire des choix fondés sur des connaissances plutôt que sur des croyances folkloriques ou des traditions.

6.2.5 Le cinquième champ de compétence: Les dimensions interpersonnelles de l'évaluation

Les variables personnelles et interpersonnelles sont sources d'influence sur les pratiques d'évaluation, de correction et de notation. Plusieurs de ces variables font pression sur les professeurs au moment d'attribuer une note à un étudiant ou à un groupe d'étudiants. Selon la nature de ces variables et, surtout, selon la conception qu'on se fait de l'évaluation et de la fonction des notes, on aura plus ou moins tendance à accepter de nuancer le jugement final que constitue la note (Stiggins, 1991).

La revue de la documentation (chapitre 2) décrit un certain nombre d'attitudes qui influencent directement les pratiques en matière d'évaluation. On y constate que les enseignants sont nombreux à chercher un équilibre entre l'indulgence et l'intransigeance et qu'ils optent, en général, pour une attitude indulgente.

Analyse

Dans le cadre de la troisième question du questionnaire, on mesure la pratique des professeurs qui choisissent a priori d'accorder des points, au bénéfice de tous les étudiants du groupe, pour des variables qui ne font généralement pas partie des résultats attendus de l'apprentissage. On sait, par exemple, que des professeurs indiquent à leur plan de cours qu'ils accorderont des points pour la présence assidue au cours.

Dans la présente étude sur les variables interpersonnelles qui jouent sur les pratiques d'évaluation, les questions 45 et 46 mesurent les proportions de répondants qui disent accepter de modifier des notes finales, non pas pour un groupe entier, mais pour des individus et ce, en fonction des comportements de ces individus.

Il y a symétrie entre la plupart des énoncés des deux groupes de pratiques d'attribution de notes, à savoir:

"J'accorde des points pour":

- la motivation et l'effort dans l'apprentissage;
- la présence assidue aux cours;
- l'exécution de devoirs ou d'exercices à la maison ou en classe;
- les progrès réalisés au cours de la session (évolution);
- la participation active de l'étudiant en classe;
- la créativité, l'originalité.

Malgré cette symétrie, nous constatons que la pratique de récompenser ou de pénaliser des individus suite à certains comportements (Q45 et Q46) ne coïncide pas nécessairement avec la pratique d'accorder des points à l'ensemble des étudiants pour les mêmes comportements (Q3).

Le tableau 4 montre les corrélations qui existent entre les énoncés équivalents de la série "Question 3" et de la série "Questions 45 et 46". On y voit que les corrélations entre les pratiques semblables dans les deux séries se situent à des niveaux modestes quoique tous positifs. C'est dans la pratique de "récompenser" la participation active des étudiants en classe que la corrélation est la plus forte (τ = .363) entre cette pratique "pour tous" (Q3) et cette pratique pour des individus, en fin de session (Q46). C'est sur l'exécution de devoirs ou exercices à la maison ou en classe que la corrélation est la plus faible (τ = .125).

Tableau 4: Corrélations entre les réponses aux variables équivalentes des questions 3 et 45, 46.

	3a) la motivation et l'effort	3b) présence assidue aux cours	3c) devoirs ou exercices à la maison ou en classe	3d) progrès réalisés au cours de la session	3e) participation active en classe	3f) la créativité, l'originalité
Q45) Diminuer pour pénaliser:						
45a) absences non motivées		,344				
45b) peu d'effort fourni	,303					
Q46) Augmenter pour récompenser:						
46a) motivation et effort	,25					
46b) présence assidue		,29				
46c) devoirs à la maison			,125			
46e) progrès réalisés				,326		
46f) participation en classe					,363	
46g) créativité, originalité						,335

Note: Q46d) "réalisation de travaux supplémentaires" n'a pas de réciproque dans la Q3. La Q3g) "le respect de mes exigences de présentation des travaux" n'a pas de réciproque dans la Q46.

Des corrélations de cet ordre signifient qu'il y a peu de relation entre, d'une part, la pratique d'attribuer des points pour certains comportements et ce, pour tous les étudiants, tel que mesuré par la question 3 au chapitre des cibles de l'évaluation et, d'autre part, la pratique d'augmenter ou de diminuer des notes individuelles, en fin de session. Bref, le professeur qui a tendance à modifier, à sa discrétion, la note finale d'un étudiant pour pénaliser ou pour récompenser certains comportements aura plutôt rarement la politique de cibler ces mêmes comportements lors de ses évaluations "pour tous" (Q3). Il s'agit donc là d'une politique discrétionnaire qui sera mise en oeuvre pour gérer des cas particuliers.

Certaines contradictions sont apparentes entre deux énoncés de croyances. On pourrait s'attendre à ce que ceux qui croient qu'il ne faut pas augmenter la note d'un étudiant pour l'encourager à travailler davantage (44c) croiront aussi que tous les étudiants d'un même cours devraient être évalués à partir des mêmes critères (33d). Pourtant, une corrélation de rang (Kendall) entre ces deux énoncés révèle une relation faible quoique positive entre ces 2 variables (tau = 0,146; p = ,0001). Par un tableau de contingence valide (p = .0001), on remarque que chez ceux qui croient que les étudiants devraient être évalués à partir des mêmes critères (33d), 25% ne croient pas qu'ils sont influencés par l'identité du répondant (44d) lors de la correction. Chez ceux qui ne savent pas si les étudiants d'un même cours devraient être évalués à partir des mêmes critères (33d), tous sont d'accord qu'il ne faut pas augmenter la note d'un étudiant pour le pousser à travailler davantage (44c).

Ces incohérences révèlent, croyons-nous, l'étendue du malaise que vivent les professeurs lorsque vient le temps d'établir les critères d'évaluation et de traiter des cas particuliers.

Analyse factorielle

Une analyse factorielle a été effectuée sur l'ensemble des variables de ce champ de compétence. Quatre facteurs sont identifiés (annexe 6).

Facteur 1: Le premier facteur regroupe les réponses aux questions de pratiques 45 et 46 et à l'énoncé de croyance 44a. Les fortes corrélations entre ces trois variables illustrent la pratique de pénaliser ou de récompenser des comportements individuels de certains étudiants, notamment sur la foi du mérite. Ainsi, les professeurs qui croient que les étudiants qui travaillent fort méritent d'avoir des notes plus élevées que ceux qui ne le font pas (44a) ont aussi tendance à diminuer une note finale pour pénaliser des absences non motivées ou des efforts insuffisants. Ce sont aussi ceux qui ont tendance à augmenter une note finale pour récompenser certains comportements, dont, principalement, les progrès réalisés pendant la session ainsi que la motivation et l'effort dans l'apprentissage. Ce premier facteur est celui autour duquel s'assemblent les professeurs qui se serviront des notes comme "bâton" ou comme "carotte".

Facteur 2: Le second facteur regroupe les professeurs qui manipulent les notes avec rigueur. Plus ils croient qu'il faut s'abstenir d'augmenter la note d'un étudiant pour l'encourager ou le pousser à travailler davantage (Q44c), plus rarement ont-ils tendance à encourager un étudiant faible en accordant une note de passage lorsque ses efforts semblent appréciables (Q47) et à modifier les notes de tous pour ajuster la moyenne de classe (Q48).

Facteur 3: Ce facteur caractérise les professeurs qui, semble-t-il, cherchent à "passer des messages" par le truchement de leurs évaluations. Plus ils croient que de mauvais résultats d'évaluation ont un effet démotivant sur les étudiants (44b), plus ils ont tendance à donner un examen plutôt facile pour

compenser un examen antérieur difficile (Q23). On y trouve les professeurs probablement plus indulgents, ceux qui vont chercher à "réparer les pots cassés" après un examen difficile.

Facteur 4: Le quatrième et dernier facteur mesure la conscience que des influences externes peuvent intervenir dans l'évaluation des apprentissages. Ce facteur est composé de deux croyances, à savoir que, lors de la correction des questions à développement ou des travaux écrits, les enseignants sont influencés par la connaissance de l'identité du répondant (44d) et que c'est la conception de l'enseignement qui influence le plus nos choix de stratégies d'évaluation (44e).

Le premier et le troisième facteur méritent une attention particulière au chapitre de la formation et du perfectionnement. D'une part, au vu du premier facteur, il faut sensibiliser les professeurs aux problèmes d'équité et d'interprétabilité des notes que soulève la pratique de moduler des notes individuelles sur la foi de variables qui sont externes aux compétences à atteindre. Ils doivent être conscients des conséquences éducatives de cette pratique. D'autre part, en fonction du troisième facteur, il faut aider les professeurs à comprendre et à gérer les jeux d'émotions qui touchent certains étudiants lorsqu'ils obtiennent des "mauvais résultats".

Groupes réacteurs

Suite à la présentation de nos statistiques auprès des groupes réacteurs, certains commentaires nous invitent à la prudence dans l'interprétation de nos chiffres. D'autres, cependant, semblent confirmer une réalité bien établie.

Sur les pratiques de récompenser ou de pénaliser des individus suite à certains comportements (Q45 et Q46), deux professeurs de mathématiques disent estimer que les pourcentages obtenus à la question 46 sont surestimés. Par exemple, cela peut être une politique annoncée au plan de cours que d'arrondir une note faible si l'étudiant a affiché des progrès

évidents ou une bonne volonté manifestée notamment au chapitre de la participation en classe. Selon ces professeurs, cette situation peut se rencontrer dans les cours de mise à niveau comme le cours de Mathématiques d'appoint 201-311.

Un autre professeur pense que des répondants pourraient avoir mal compris la question et y avoir répondu comme s'il s'agissait de la question 3, qui explore les pratiques de notation sur des cibles autres que l'apprentissage.

Plusieurs autres commentaires apportent des exemples de situations qui seraient de nature à favoriser la pratique de récompenser ou pénaliser des individus suite à certains comportements. Par exemple, un professeur affirme qu'il peut être légitime d'ajuster à la hausse ou de donner une note de passage à la fin de la session, si les progrès réalisés par un étudiant sont tels que le professeur a réellement l'impression d'avoir sous-estimé cet étudiant jusqu'à ce moment-là.

Un professeur de chimie dit qu'il peut tenir compte de la motivation et l'effort, de la participation en classe, etc., pour juger de l'opportunité de "faire passer" un étudiant qui n'aurait que 57% à la fin de la session. Il se sert de ces considérations pour les quelques étudiants qui, en fin de session, se situent juste en-dessous du seuil de réussite, cette zone grise qui se situe pour certains professeurs entre 57% et 60%, et pour d'autres, entre 55% et 60%. L'étudiant qui aura manifesté sa motivation par des efforts apparents ou par une participation active pourra amener le professeur à donner la chance au coureur et à lui accorder l'indulgence de la note de passage.

Un professeur appartenant à la strate Arts (pré-universitaire et technique) affirme que toutes les variables proposées à la question 46 sont incluses de façon habituelle comme critères de notation dans son département. Le tableau C5-b, en annexe 5, semble confirmer ce commentaire puisque, outre le groupe "Éducation physique", les proportions les plus fortes de "oui" se trouvent souvent en Arts.

Des professeurs ajoutent qu'il leur arrive aussi de considérer l'effort ou la qualité de participation pour accorder une note de passage, non pas seulement à la fin de la session, mais en cours de session, si un résultat d'évaluation devait être particulièrement démotivant pour un étudiant. Un professeur considère même que ça peut faire une différence de donner 80 plutôt que 79 à un étudiant, en cours de session, et que, pour poser ce geste, il tiendra compte des variables proposées dans la question 46.

En Techniques biologiques, une enseignante affirme que ces dimensions très subjectives peuvent intervenir de façon incontournable dans le cadre des stages. La relation qui se développe entre une enseignante et un petit nombre d'étudiants stagiaires peut mettre en évidence des caractéristiques personnelles et favoriser que ces caractéristiques interviennent lors de l'attribution de la note. C'est pourquoi, dit-elle, il n'est pas rare qu'un comité de révision de note en arrive à juger qu'il est justifié d'invoquer les efforts d'une étudiante pour lui accorder la note de passage. Ce phénomène est assurément présent, selon cette enseignante, dans les réponses obtenues à l'énoncé de pratique (Q47): "J'ai tendance à encourager un étudiant faible en accordant une note de passage lorsque ses efforts me semblent appréciables".

À ces considérations s'ajoute peut-être, selon un enseignant, une réalité plutôt organisationnelle. En effet, il invoque les arguments suivants pour justifier la pratique de pénaliser ou de récompenser certains comportements d'étudiants. Selon lui,

- de plus en plus d'étudiants, dans les salles de classe, sont moins disposés à apprendre. Leur motivation ou leur capacité d'attention est amoindrie par divers facteurs: la fatigue à cause du travail rémunéré, l'éclatement des valeurs, les séparations familiales, l'hétérogénéité des préalables académiques;

- ces variables entraînent une augmentation des problèmes de discipline en classe et les professeurs consacrent, plus que jamais, du temps à gérer la classe, à assurer la discipline. Or, les cégeps ne disposent pas de politique disciplinaire et les

professeurs sont laissés à eux-mêmes pour faire face à une mosaïque comportementale complexe, particulièrement en première année de cégep;

- en conséquence, plusieurs professeurs cèdent à la tentation d'utiliser la note comme "bâton" ou comme "carotte" parce que c'est le seul moyen efficace de contrôler la classe.

Interprétation

Les réponses obtenues aux questions 23 et 48 donnent à penser que certains professeurs éprouvent des difficultés à planifier ou à gérer leurs évaluations des apprentissages de leurs étudiants. Il arrive que certains examens s'avèrent difficiles, que la moyenne de la classe est basse, que les échecs soient nombreux. Pourtant, certains professeurs se comportent comme si c'était un accident. Peut-être ont-ils mal diagnostiqué les préalables et les besoins de leurs étudiants, en début de session. Peut-être ont-ils mal ajusté l'examen au contenu ciblé, à l'aide par exemple d'un tableau de spécification. Peut-être enfin ont-ils posé des questions trop difficiles, n'ayant pas fait l'analyse statistique des items. S'il arrive que certains professeurs donnent un "examen plutôt facile" (Q23) ou cherchent à "ajuster la moyenne de classe" (Q48) en réaction à des évaluations accidentellement trop difficiles, d'autres enseignants utilisent leurs évaluations en les calibrant de façon à provoquer certains effets relatifs à la gestion de classe. En effet, certains professeurs semblent doser le degré de difficulté de leurs évaluations, expressément pour motiver ou pour faire travailler les étudiants. On fait un examen difficile pour passer un message. On en fait un autre, plus facile, pour passer un autre message.

Nos données confirment les observations de Réginald Grégoire (Conseil des collèges, 1985) à l'effet que les professeurs cherchent souvent à donner une chance à l'étudiant et, en conséquence, adoptent une attitude plutôt indulgente. Les professeurs savent qu'une "mauvaise note" peut avoir pour effet de décourager des étudiants. Selon les circonstances, la définition qu'on se fait d'une "mauvaise note" peut varier. Pour certains

étudiants, une note de 79% peut être aussi décourageante que ne l'est un 58% pour d'autres. En cours de session, un échec à 40% peut paraître plus insurmontable qu'un échec à 55%. C'est pourquoi les professeurs cherchent souvent à atténuer l'effet de démotivation d'une mauvaise note en tenant compte des cas particuliers et de la sensibilité de certains étudiants.

Il est évident que plusieurs facteurs peuvent intervenir sur les décisions que les professeurs prennent lorsque vient le temps de gérer les notes, surtout les notes finales. Si certaines pratiques peuvent se justifier selon les contextes, d'autres posent certainement des problèmes d'équité, voire d'éthique.

Le concept de "devoirs ou travaux à la maison ou en classe" nécessite des clarifications. La pratique de donner des points pour l'exécution de devoirs ou travaux mérite assurément d'être discutée selon la compréhension qu'on s'en fait. Un devoir à la maison est-il un exercice d'apprentissage ou est-il un processus de travail qui fait l'objet d'une évaluation? Si c'est une "cible" d'évaluation, parlerons-nous d'évaluation formative ou sommative? Si on augmente une note finale d'un étudiant sur la foi d'un devoir à la maison ou d'un travail supplémentaire, le professeur en fait-il l'équivalent d'un examen de reprise? Comment contrôle-t-il l'équivalence entre le devoir ou travail supplémentaire de cet étudiant et ce que les autres étudiants ont fait en guise d'évaluation sommative?

On ne peut pas porter de jugement de valeur sur la foi des résultats statistiques seulement, car la réalité spécifique à certaines disciplines peut justifier des nuances. Par exemple, en Éducation physique, la situation pédagogique est telle qu'on ne peut pas nécessairement juger comme incorrecte la pratique de diminuer une note finale pour pénaliser des absences non motivées ou des efforts insuffisants.

Par ailleurs, certaines politiques institutionnelles d'évaluation des apprentissages prévoient que les enseignants peuvent donner un verdict d'échec si les absences dépassent, par exemple, 15% du temps de classe. Un tel article peut être énoncé et appliqué différemment d'un collège à l'autre et il nous semble possible que certains répondants à notre

questionnaire pourraient avoir interprété notre question 45 de façon à tenir compte de leur politique institutionnelle ou départementale.

Recommandations

Recommandation 10: La formation des professeurs en mesure et évaluation devrait leur permettre de faire face aux nombreuses variables personnelles et interpersonnelles qui interviennent sur le processus d'évaluation, de correction et d'attribution des notes afin qu'ils puissent gérer ces processus de façon équitable et pédagogiquement cohérente.

Recommandation 11: Les professeurs devraient pouvoir identifier et évaluer les conséquences éducatives et sociales des diverses pratiques de notation, particulièrement en ce qui a trait à l'attribution de notes pour l'effort, la motivation, l'exécution d'exercices en classe ou hors de la classe.

Recommandation 12: Les professeurs devraient s'assurer que leurs pratiques d'évaluation et de notation soient cohérentes avec les objectifs pédagogiques et les critères annoncés dans leurs plans de cours.

Recommandation 13: Nous recommandons qu'une étude soit faite sur le degré de relation qui pourrait exister entre, d'une part, l'hétérogénéité motivationnelle des étudiants et l'absence de politique disciplinaire dans les collèges et, d'autre part, les pratiques d'augmenter ou de diminuer des notes pour récompenser ou pénaliser certains comportements.

6.2.6 Le sixième champ de compétence: La rétroaction et la notation

Analyse

Les résultats aux énoncés appartenant à ce champ de compétence révèlent des malaises, chez les professeurs, à l'égard de leur propre subjectivité en évaluation (Q49f), de l'approche normative de l'évaluation (Q49b), du rôle que jouent les notes en éducation (Q49h et Q49c). Ces résultats donnent à penser que la conception que plusieurs professeurs ont de l'évaluation serait tributaire de l'ancien paradigme de l'évaluation, décrit au chapitre trois. Plusieurs croyances et pratiques illustrent la méconnaissance des principes de mesure et évaluation et fondent probablement ce malaise de nombreux professeurs à l'égard de l'évaluation.

Un grand nombre croient que la rétroaction exige des pratiques qui, pourtant, augmentent considérablement le fardeau de la correction (Q49e et Q49i), par exemple, la rédaction de longs commentaires sur les copies.

Les pratiques d'attribution des notes dans le cadre du travail en équipe suggèrent que plusieurs s'exposent à ce que les notes des étudiants ne correspondent pas nécessairement à la compétence personnelle de chacun.

Les pratiques de notation en évaluation formative montrent clairement que deux professeurs sur cinq pratiquent une évaluation formative qui n'en est pas. La pratique de verser au bulletin de l'étudiant des notes obtenues dans des évaluations formatives (Q50) n'est cependant pas liée à la croyance très répandue qui veut que la note, ça représente la paye de l'étudiant (Q49h), car il n'y a pas de corrélation entre ces deux variables.

L'effet des moyennes de la classe

L'une de nos questions cherche à savoir si les professeurs croient que la note moyenne d'un groupe est un indicateur de la qualité de l'enseignement. Combien de professeurs croient que cette relation est vraie? Combien n'y croient pas? Un groupe cours peut-il afficher une moyenne

élevée sans que la qualité y soit pour quelque chose? Les pratiques de notation d'un professeur peuvent-elles générer une moyenne de groupe qui ne soit pas un reflet direct de la qualité de l'enseignement reçu? En groupes réacteurs, l'interprétation dominante des résultats observés par le questionnaire est à l'effet que les moyennes peuvent varier sans égard à la qualité de l'enseignement.

La tendance majoritaire d'un groupe réacteur est de lire cette question telle que formulée, à la lettre, c'est-à-dire que 76% des professeurs répondants pensent qu'il n'y a pas de relation entre la variable "qualité de l'enseignement" et la moyenne de la classe, telle qu'observée au bulletin final. Comme si ces professeurs nous disaient, ici, que la note moyenne de la classe ne varie pas en fonction de la qualité de l'enseignement. Plusieurs professeurs, dans la salle, ajoutent que, selon eux, à qualité égale d'enseignement, les moyennes des groupes peuvent varier beaucoup à cause des caractéristiques des clientèles. L'homogénéité ou l'hétérogénéité sont des variables qui peuvent faire fluctuer les moyennes de classe. Mais est-ce la seule explication, la seule piste qui puisse expliquer que les moyennes de classe varient?

Les réponses à notre questionnaire nous apprennent que près d'un professeur sur cinq croit que la moyenne de sa classe est le reflet de la qualité de son enseignement. Les moyennes de classe sont des données qui permettent les comparaisons entre les professeurs et, pour certains, cela peut devenir anxiogène. On pourra s'inquiéter de la pression que peut engendrer cette croyance sur les pratiques de notation de certains professeurs lorsque la moyenne risque d'être "anormalement" faible ou élevée. Comme tout le monde, les professeurs se préoccupent du jugement d'autrui. Ils voudront faire bonne figure s'ils pensent qu'on peut les évaluer en utilisant les moyennes comme indicateurs. Des moyennes de classe plus faibles que celles des collègues du département peuvent-elles leur valoir d'être vus comme des professeurs trop sévères? Une moyenne plus élevée que celle des autres qui enseignent le même cours peut-elle valoir à ce professeur d'être jugé trop généreux, voire laxiste? Et comment les professeurs perçoivent-ils le jugement que portent les étudiants devant les moyennes de groupe?

Tous n'ont pas nécessairement cette préoccupation. Pour une enseignante d'un groupe réacteur, la moyenne est un indicateur qui peut l'aider à ajuster ses approches pédagogiques, ses stratégies.

Nos données quantitatives de même que les commentaires des groupes réacteurs nous apprennent que des professeurs s'inquiètent des conséquences de la perception qu'on peut avoir de la moyenne de classe. Mais d'autres s'en servent pour réguler leur propre pédagogie et considèrent que la moyenne constitue une information qui peut leur être utile.

Des notes et du mérite

L'énoncé 49d est destiné à évaluer l'opinion que se font les professeurs de la qualité générale du processus d'évaluation, tel qu'ils le perçoivent dans leur environnement. Rappelons-en la formulation: "Plusieurs enseignants accordent une note de passage à des étudiants qui ne la méritent pas."

Dans les réponses obtenues, le taux de "ne sais pas" est extrêmement élevé (41%), soit parce que les répondants ne savent vraiment pas ce qu'il en est dans leur entourage, soit parce qu'ils choisissent de ne pas se prononcer, pour des raisons éthiques évidentes. Mais chez ceux qui expriment leur opinion, l'accord avec cet énoncé est légèrement plus élevé (32%) que le désaccord (27%).

Dans les groupes réacteurs, des professeurs pensent que certains répondants ont peut-être pensé, ici, aux situations lors desquelles un étudiant a une note finale de 58% ou 59% qu'on "arrondit" à 60% à partir de la notion de "doute raisonnable". Le rapport Grégoire (1985) faisait allusion à cette pratique en disant que beaucoup de professeurs, devant une note finale juste en deçà de la note de passage, adoptent une attitude d'indulgence.

Dans un environnement de recherche analogue au nôtre, les chercheurs Crowl et Berkowitz (1985) ont rapporté que les professeurs interrogés dans

leur enquête se disent très insatisfaits des pratiques et des standards d'évaluation dans leurs écoles. Afin d'interpréter cette insatisfaction, ces chercheurs ont fait l'hypothèse, entre autres, que leurs répondants réagissaient peut-être à la forte pression morale qui pourrait s'exercer sur eux à l'effet de "faire passer" certains étudiants qui devraient pourtant échouer. Cette hypothèse vaut la peine d'être rappelée, ici, lorsque nous considérons les résultats observés à notre question 49d). Est-il possible que certains de nos répondants expriment, ici, qu'ils se sentent parfois contraints de "faire passer", à regret, des étudiants qui ne le méritent pas et ce, sous une quelconque pression morale venant de leur entourage? Cette possibilité vaudrait la peine d'être explorée plus à fond car il est notoire, dans certains milieux, que l'emploi de professeurs dépend du nombre d'étudiants qui réussissent des cours préalables. Par exemple, le nombre de groupes-cours en deuxième année est fonction du nombre d'étudiants qui réussissent les cours de première année. Cette considération a une incidence directe sur l'embauche ou la mise en disponibilité des professeurs. Pourrait-elle constituer, en soi, une pression morale qui incite des enseignants à relâcher un peu au chapitre de la rigueur?

En somme, cet énoncé de croyances nous rapporte des réponses qui constituent, croyons-nous, un autre jugement sévère sur les pratiques d'évaluation telles qu'elles sont perçues dans l'entourage des répondants. Rappelons que les résultats sont semblables pour l'énoncé 33h qui mesure, dans le troisième champ de compétence, la perception de rigueur dans les pratiques des autres professeurs.

Analyse factorielle

Les variables comprises dans le sixième champ de compétence sont soumises à l'analyse factorielle qui extrait sept facteurs. Mais les trois derniers facteurs sont composés d'associations que nous qualifions de fortuites. Ces facteurs n'ont pas de logique interne et nous les ignorons, ci-après, car ils n'ont pas de signification utile. Les quatre premiers facteurs sont:

Facteur 1: La "gratuité" de l'évaluation formative. Le premier facteur est celui qui reconnait que l'évaluation formative constitue une "zone franche" dans laquelle l'étudiant peut apprendre de ses erreurs sans risque de pénalité. La pratique de verser au bulletin final des notes obtenues en évaluation formative (Q50) est liée très fortement à la croyance que l'évaluation formative doit ou ne doit pas compter au bulletin (Q49g).

Facteur 2: Les rétroactions correctives. Ce facteur caractérise les pratiques de rétroaction correctives. C'est sous ce facteur qu'on valorise le feedback (Q49e) et qu'on donne aux étudiants les moyens d'améliorer leurs points faibles.

Facteur 3: L'évaluation normative. Ce facteur décrit la conception selon laquelle chaque étudiant doit avoir sa propre note (Q28) et que la moyenne et la dispersion des notes autour de la moyenne sont des variables comparatives importantes (Q49a et b). C'est la conception normative de l'évaluation.

Facteur 4: La note comme récompense. Ce facteur résume la conception selon laquelle les étudiants valorisent la note et les commentaires manuscrits de leurs professeurs comme des récompenses ou comme des indicateurs du mérite de la production étudiante. Selon ce facteur, les copies d'examens et les travaux écrits sont vus comme des objets qui sont échangés pour des notes ou des commentaires appréciatifs, un peu comme dans une situation de troc.

Nous avons vu que les professeurs utilisent plusieurs approches d'évaluation des apprentissages, visent diverses cibles d'apprentissage et incorporent souvent à la note finale des notes obtenues de l'observation de diverses variables extérieures à la discipline (motivation, participation, etc.)

Mais nous ne savons pas comment les professeurs combinent les mesures (scores) de diverses sources pour aboutir à calculer une note finale. Nous ne savons pas si les diverses sources de données sont pondérées entre elles. Par exemple, comment arrivent-ils à décider d'utiliser deux examens objectifs, un travail de session, un journal de bord et un travail d'équipe? Comment en arrivent-ils à pondérer ces dispositifs d'évaluation pour déterminer que les examens vaudront 15% chacun? Certes, les plans de cours nous informent de la distribution des points accordés à chaque dispositif mais nous ne savons pas si ces points sont attribués arbitrairement ou si les professeurs mettent à contribution une méthodologie de planification avec définition de domaine, tableau de spécification, taxonomie d'objectifs, pondération d'objectifs.

Recommandations

Recommandation 14: Il y a lieu d'étudier le processus de décision qui mène les professeurs à combiner des évaluations de diverses nature, tant quantitatives que qualitatives, pour générer une note finale. Cette étude devrait mettre en lumière les structures de planification que les professeurs mettent en oeuvre pour pondérer les diverses mesures en fonction de leur importance relative et devrait guider la formation des professeurs en regard de cette dimension de la planification de l'évaluation. Cette étude devrait caractériser les croyances et les conceptions des professeurs à l'égard, notamment, des fonctions de l'évaluation dont particulièrement de l'évaluation formative, ainsi que du rôle des notes.

Recommandation 15: Nous recommandons que les professeurs des cégeps aient accès à une formation qui leur permette de clarifier les fonctions de la rétroaction et de la notation de façon à ce que les pratiques correspondent aux recommandations admises en mesure et évaluation. Cette formation devrait conférer aux pratiques des professeurs toute la crédibilité qu'elles doivent avoir de façon à ce que les notes reflètent fidèlement les résultats des apprentissages des étudiants.

6.3 La formation en mesure et évaluation: Analyse et recommandations

6.3.1 La préparation des enseignants en mesure et évaluation des apprentissages

L'analyse des données obtenues à la question 55 nous montrent que 35% des professeurs disent avoir suivi "au moins un cours" dans une université, au baccalauréat ou à la maîtrise, sur le thème de l'évaluation des apprentissages. Lorsqu'on ajoute ceux qui ont aussi suivi "au moins un cours" dans le cadre du programme CPEC - Performa, donc en cours d'emploi, la proportion de professeurs de cégeps ayant suivi un cours ou plus en évaluation des apprentissages, soit au baccalauréat, soit en maîtrise, soit au certificat Performa, passe à 49%. Rappelons que, si l'on peut présumer que les cours au baccalauréat ou à la maîtrise sont de trois crédits (45 heures), les cours offerts au CPEC - Performa peuvent être d'aussi peu que d'un crédit.

En plus, 15% des professeurs ont suivi "au moins un cours" dans plus d'un niveau universitaire, par exemple, un cours de baccalauréat et un cours de CPEC - Performa ou toute autre combinaison bacc. - maîtrise - CPEC. Notre questionnaire ne nous permet pas de savoir si les professeurs ont suivi plus d'un cours, si la formation reçue était pertinente aux fonctions qu'ils occupent au collégial ou si les cours du CPEC - Performa étaient d'un ou deux ou de trois crédits.

Quant aux autres répondants, l'analyse de la question 55 nous montre que 35% des professeurs des cégeps n'ont reçu aucune préparation de niveau universitaire en mesure et évaluation.

Un professeur, dans un groupe réacteur, propose une distinction pouvant expliquer les différences de formation en pédagogie ou en évaluation des apprentissages. Selon cette personne, les professeurs qui vont suivre des cours en pédagogie le font parce qu'ils sont convaincus que cela peut les aider. Ils accordent une certaine importance à la pédagogie. Par ailleurs,

les professeurs qui ne suivent aucun cours de perfectionnement en pédagogie seraient ceux pour qui la pédagogie est un art qui ne s'apprend pas: ça s'improvise avec le temps, avec l'expérience. Cet enseignant pense qu'actuellement, dans le réseau collégial, on a affaire à ces deux tendances et la tendance majoritaire serait cel!e pour qui la pédagogie, c'est un art qui ne s'enseigne pas mais qui s'apprend avec l'expérience.

6.3.2 L'autoévaluation des compétences en évaluation des apprentissages

La question 56 permet aux répondants d'autoévaluer leur compétence en évaluation, tant formative que sommative, des apprentissages. On a vu plus haut, que la majorité des professeurs jugent que leur compétence est plutôt bonne, sinon très bonne en évaluation formative (67%) de même qu'en évaluation sommative (88%). Très peu d'enseignants jugent que leur compétence est faible. À l'égard de l'évaluation sommative, personne ne juge sa compétence comme nulle.

Quelle est la valeur de l'autoévaluation des professeurs quant à leur compétence? Nous avons testé le mérite des réponses à cette question en faisant des tableaux croisés entre, d'une part, les réponses à la question 56b: "Ma compétence en évaluation sommative" et quelques exemples d'énoncés qui mesurent des croyances ou des pratiques relatives aux qualités de l'évaluation. Un tableau de contingence avec l'énoncé 33d (τ = ,095; p = ,0007) nous montre que chez ceux qui ne croient pas que tous les étudiants devraient être évalués à partir des mêmes critères, 71% qualifient de "bonne" leur compétence en évaluation sommative. De même, 87% de ceux qui "ne savent pas", sur cette même question portant sur les critères, jugent leur compétence "bonne" et "très bonne". À la question 8, on a demandé si les répondants se font un plan détaillé de ce qu'ils veulent évaluer lorsqu'ils rédigent un instrument de mesure: 93% de ceux qui ne le font "jamais" et 90% de ceux qui le font rarement ont qualifié de "bonne" ou "très bonne" leur compétence en évaluation sommative (τ = ,126; p = ,0001).

Des participants aux groupes réacteurs estiment que l'autoévaluation des compétences en évaluation sommative est crédible puisque les professeurs de cégeps ont acquis une expérience pratique en évaluation. Pour ces professeurs, les chiffres observés sont réalistes quant à la compétence en évaluation sommative. Par exemple, un professeur interprète l'évaluation favorable que les répondants ont fait de leur compétence en évaluation formative et en évaluation sommative par le constat de l'âge moyen des professeurs du réseau. Le niveau d'âge est élevé et les professeurs se sentent peut-être suffisamment expérimentés en pédagogie et en évaluation. D'autres expriment un doute sur la crédibilité de cette autoévaluation, disant que s'ils font profession d'enseigner et s'ils affirment que l'évaluation est une partie intégrante de leur profession, il leur serait impossible d'admettre une quelconque incompétence en évaluation car cela serait l'équivalent d'admettre une incompétence à enseigner. Sur l'évaluation formative, cependant, plusieurs participants pensent que certains répondants auront surestimé leur compétence, d'autant plus que, puisqu'il a été évident que le concept d'évaluation formative est mal connu, il est facile de mal juger une compétence dans un domaine qu'on connaît mal.

On observe dans cette recherche que les concepts de base en évaluation sont souvent méconnus. Par exemple, on confond l'évaluation sommative et formative, habiletés, aptitudes, objectifs et compétences. Une enseignante qui a 10 ans d'expérience écrivait: "... il ne faut pas mesurer seulement le objectifs mais aussi les habiletés". Cette confusion dans le vocabulaire est vraisemblablement le fruit d'une culture pédagogique folklorique acquise "sur le tas" dans les collèges et dans les départements. Ces observations donnent à penser qu'un certain nombre de professeurs de cégeps ne connaissent pas leur incompétence en évaluation des apprentissages. Selon un participant dans un groupe réacteur, "moins on en sait sur l'évaluation, moins on peut s'apercevoir qu'il nous manque quelque chose. On ne peut pas se trouver incorrects si on ne sait pas qu'il y a autre chose".

Au fil des années, les professeurs acquièrent une expérience pratique qui leur permet de résoudre la plupart des problèmes qu'ils rencontrent en évaluation. Ainsi, d'une façon ou d'une autre, beaucoup de professeurs apprennent à effectuer leur tâche en s'ajustant, sur le terrain, aux divers problèmes et ce, même en l'absence de quelque formation en mesure et évaluation. Cela peut vraisemblablement générer un sentiment d'équilibre, voire un sentiment de compétence qui pourrait avoir contribué à ce que certains jugent leur compétence comme bonne et même très bonne.

Au vu des données recueillies ainsi que des analyses que nous en faisons, nous proposons, ci-après, des recommandations concernant la formation des professeurs, la formation des cadres et des conseillers pédagogiques des collèges, les collaborations souhaitables en matière de développement de compétences, les services à offrir.

6.3.3 La formation des professeurs

Plusieurs chercheurs ont observé que les pratiques de mesure et d'évaluation des apprentissages sont très diversifiées, même à l'intérieur des groupes disciplinaires, que les attitudes et les croyances à l'égard de l'évaluation et de la notation sont aussi très variées, et, enfin, qu'il y a persistance de certaines fausses conceptions au sujet de la mesure, de l'évaluation et de la notation (Boothroyd, 1992; Fontaine, 1988; Frary, 1992; Jett et Schafer, 1992; Marso et Pigge, 1992; Nadeau et Comeau, 1975; Schafer, 1991; Stiggins, 1992).

Dans la présente recherche, nous faisons les mêmes constats et nous nous joignons à tous ces auteurs qui concluent qu'il faut rehausser le niveau de la formation des intervenants en évaluation des apprentissages dont celle, tout particulièrement, des professeurs. Les résultats de notre recherche nous incitent à formuler des recommandations touchant les cibles, les lieux et la nature de la formation qui devrait être proposée aux professeurs des cégeps.

Les cibles de la formation

Chacun des 6 champs de compétence de Stiggins

Les six catégories de compétence que nous avons utilisées dans la présente recherche se sont avérées complètes, cohérentes et suffisamment simples pour convenir à notre désir de circonscrire le domaine des croyances et des pratiques en évaluation des apprentissages au collégial. Ces six catégories ou champs de compétence ont été jugés valides et pertinents par les juges experts que nous avons consultés lors de la validation de notre questionnaire de recherche. Enfin, nous avons utilisé avec satisfaction ces mêmes six catégories de compétence lors de la présentation de nos résultats de recherche auprès des groupes réacteurs. Les six champs de compétence décrits par Stiggins nous apparaissent bien adaptés à la réalité de l'évaluation des apprentissages en salle de classe au collégial.

Recommandation

Recommandation 16: Nous recommandons que les six champs de compétence décrits par Stiggins et repris ci-après, en annexe 7, servent à structurer la formation et le perfectionnement des professeurs de cégeps en matière de mesure et évaluation des apprentissages.

Les cibles particulières dans chacun des 6 champs

La lecture des résultats de recherche, présentés au chapitre 5 et dans les analyses qui précèdent, révèle des lacunes particulières, chez plusieurs enseignants de cégeps. On a noté une méconnaissance assez importante du concept de l'évaluation formative. Par ailleurs, les notes qui sont versées aux bulletins des étudiants semblent souvent polluées par l'introduction de variables qui n'ont rien à voir avec l'atteinte des objectifs d'apprentissage. Plusieurs professeurs font du détournement de

l'évaluation en s'en servant à diverses fins, plus ou moins pertinentes aux véritables fonctions de l'évaluation des apprentissages en salle de classe. Enfin, nous avons noté la présence de conceptions de l'évaluation qui nous semblent pernicieuses et qui génèrent probablement des attitudes hostiles, lorsqu'on questionne les pratiques des professeurs ou des attitudes qui n'incitent pas certains à se préoccuper de développer des compétence en évaluation des apprentissages.

Recommandation

Recommandation 17: Nous recommandons que des efforts de formation soient engagés, en priorité, pour corriger les lacunes les plus importantes. Ces lacunes, que nous considérons prioritaires, sont présentées, ci-après, en regard de chacun des six champs de compétence décrits par Stiggins.

Les 6 champs de compétence de Stiggins	Les cibles prioritaires de la formation
1. L'utilisation de l'évaluation en classe.	- Le détournement de l'évaluation (se servir de l'évaluation pour gérer la classe plutôt que pour aider l'apprentissage); - L'évaluation formative.
2. Les cibles de l'évaluation.	- La pollution de la note des groupes (la carotte pour faire travailler).
3. Les qualités de l'évaluation.	- La planification et la validation des dispositifs d'évaluation (l'isolement des professeurs et la peur d'être jugés: on ne sollicite pas des collègues pour valider ce qu'on fait).
4. Les instruments d'évaluation.	- Le besoin de diversifier les dispositifs d'évaluation.
5. Les dimensions interpersonnelles de l'évaluation.	- La pollution de la note des individus; - La normalisation informelle des moyennes de classe.
6. La rétroaction et la notation.	- La conception de l'évaluation (ancien ou nouveau paradigme); - La conception de la note; - La crédibilité de la note publiée au bulletin.

Des activités de perfectionnement en mesure et évaluation des apprentissages

En cohérence avec nos objectifs de recherche, nous avons introduit, dans notre questionnaire, une question qui vise à mesurer l'intérêt des répondants envers certains thèmes de perfectionnement. En effet, la question 57 se lit comme suit: "Si on organisait des activités de perfectionnement en mesure et évaluation, quels sujets, parmi les suivants, vous intéresseraient davantage?"

Recommandation

Recommandation 18: La dernière partie du chapitre 5 décrit l'intérêt des répondants envers certains thèmes de perfectionnement et nous recommandons que les sujets les plus demandés fassent l'objet, en toute priorité, d'offres d'activités de perfectionnement dans les collèges.

Tenir compte des croyances d'abord

Dans sa recherche, France Fontaine avait déjà observé que les croyances sont en amont des pratiques et que toute démarche visant à modifier les pratiques doit tenir compte des croyances (Fontaine, 1988, p. 5). Nos données de recherche confirment cette équation: toute entreprise d'intervention sur les pratiques d'évaluation des apprentissages doit être fondée sur une bonne connaissance des croyances des professeurs à l'égard de l'évaluation.

Recommandation

Recommandation 19: C'est pourquoi nous recommandons que la formation des professeurs tienne compte des croyances sous-jacentes aux pratiques et visent, prioritairement, la modification de ces croyances. À cet effet, les curriculums de formation devraient prévoir un diagnostic des croyances présentes et une identification de profils de croyances.

Nous croyons qu'il est probable que des combinaisons de croyances donnent lieu à des profils qu'on appelle souvent des "conceptions de l'enseignement" ou des "conceptions de l'évaluation". Nous ne savons pas, dans cette recherche, comment se jouent les relations entre les conceptions de l'enseignement, les conceptions de l'évaluation et les croyances à l'égard de l'évaluation. Nous admettons que ceux qui croient que la note est la paye de l'étudiant participent probablement à une conception de l'évaluation, différente de ceux qui ne croient pas que la note donnée à l'étudiant constitue, en quelque sorte, une paye. Il y a donc, certainement, une piste de recherche future dans la question des relations entre croyances, conceptions de l'enseignement ou de l'évaluation et pratiques en évaluation. Quelle est la contribution des conceptions de l'enseignement aux croyances dont on sait qu'elles sont en amont des pratiques évaluatives?

D'ailleurs, dans nos entrevues de pré-consultation, une enseignante nous a affirmé que les programmes de perfectionnement des professeurs et les contenus des cours en mesure et évaluation devraient être construits en fonction des conceptions de l'évaluation. C'est là que se jouent les croyances et c'est dans ce cadre qu'il faut adopter des stratégies différenciées pour faire évoluer les pratiques. On aura des stratégies de formation différentes selon qu'on s'adresse à des professeurs qui valorisent l'évaluation ou à des professeurs qui voient l'évaluation comme un fardeau agaçant.

Nous avons effectué une analyse factorielle afin de brosser un tableau des types de conceptions de l'évaluation. Sur l'ensemble des variables

produites par les réponses à notre questionnaire, nous avons choisi celles qui, selon nous, contribuent le plus à témoigner des conceptions que les professeurs se font de l'évaluation. Les variables choisies sont extraites des six champs de compétence, sans égard au libellé de chaque champ. Nous avons choisi 23 variables mesurant soit des croyances, soit des pratiques, et l'analyse factorielle a produit les six facteurs suivants:

Facteur 1: La note, un outil behavioriste. Le premier facteur caractérise la pratique des professeurs qui acceptent de moduler la note finale d'un étudiant en fonction du comportement de ce dernier. Ce facteur associe fortement la tendance à encourager un étudiant faible, s'il a fait des efforts (Q47), à la pratique de récompenser ou de punir certains comportements (Q45 et 46), en jouant à la hausse ou à la baisse sur la note finale. Ainsi, ce facteur caractérise le professeur "behavioriste" et, puisque la pratique de punir en baissant la note finale est plutôt minoritaire, par rapport à la pratique de récompenser en augmentant la note, on pourrait, à la rigueur, nommer ce facteur sous l'expression de "behavioriste indulgent".

Facteur 2: La négation de l'évaluation. Ce second facteur contient trois croyances (Q6f, 6k et 6j), toutes caractéristiques d'une conception négative de l'évaluation. Ce facteur décrit ceux qui croient que l'évaluation n'est pas utile, ni à leur enseignement, ni aux apprentissages de leurs étudiants.

Facteur 3: La normalisation des notes. Ce troisième facteur décrit la conception qui mène des professeurs à chercher à normaliser des moyennes de classe, soit en modifiant les notes des étudiants (Q48), soit en ajustant le calibre d'un examen (Q23).

Facteur 4: L'évaluation sans stress. Ce facteur décrit les enseignants qui se préoccupent des effets démotivants que peuvent avoir, sur les étudiants, les résultats de l'évaluation (Q44b). Ce sont ces professeurs qui accepteront de soumettre des questions à développement auxquels les étudiants répondront à la maison (Q25) ou qui croient qu'il faut être prêt à réajuster le contenu d'un cours en fonction des résultats des évaluation (Q6d).

Facteur 5: La note, il faut la mériter. Ce facteur est caractérisé par deux croyances centrées sur le mérite. Il décrit les professeurs qui croient que les étudiants doivent obtenir une note en fonction de leurs apprentissages réels et non en fonction des efforts qu'ils font (Q49d et 44a).

Facteur 6: Une notation rigoureuse, centrée sur l'apprentissage. Ce dernier facteur décrit la conception selon laquelle la notation est inhérente au processus d'enseignement et d'apprentissage (Q49c), et qu'en conséquence, il ne faut pas polluer la note à des fins de contrôle comportemental (Q44c), car la note doit être produite avec rigueur (Q33h).

Recommandation

Recommandation 20: Les curriculums de formation ainsi que toute production de matériel didactique, associé au développement des compétences en mesure et évaluation, devraient être fondés sur un diagnostic et une identification des profils de croyances ainsi que des conceptions de l'évaluation, afin de différencier les efforts de formation pour les adapter aux profils des clientèles-cibles.

Les lieux de la formation

Formation dans les départements

Dans les départements, plusieurs professeurs mettent au point leurs dispositifs d'évaluation dans la solitude. Cela a fait l'objet de commentaires dans les groupes réacteurs: on s'isole car on a peur du jugement des autres. "Parfois, les autres ne pensent pas comme nous. C'est pourquoi on n'échange pas". Selon des enseignants, des projets de formation intra-départementaux mettraient les professeurs en situation d'échange et favoriseraient un dialogue constructif.

Recommandation

Recommandation 21: Puisque les départements disciplinaires sont vraisemblablement la source de la formation folklorique des professeurs en matière d'évaluation et que des profils de pratiques spécifiques peuvent être identifiés en fonction des disciplines, nous recommandons que l'on offre des profils de formation qui soient adaptés aux besoins spécifiques des disciplines ou des départements.

Formation inter-départementale

Plusieurs besoins de formation en mesure et évaluation débordent du cadre du département ou du programme. Par exemple, dans les programmes du secteur technique, les laboratoires et les stages posent des problèmes de mesure et d'évaluation qui sont particuliers. De plus, l'inclusion d'objectifs pédagogiques, relevant de ce qu'on appelle encore la formation fondamentale, pose le problème de la mesure et de l'évaluation de ce type de compétence.

Recommandation

Recommandation 22: Les professeurs des cégeps devraient développer et maintenir leurs compétences en mesure et évaluation, tant dans des environnements de laboratoires ou de stages en milieu de travail, que dans le contexte du développement de compétences dites fondamentales ou générales.

Jumelages de nouveaux professeurs et d'anciens professeurs

Stiggins et Bridgeford (1985) déplorent le manque de collaboration entre enseignants, au chapitre de l'évaluation des apprentissages, alors que les recherches démontrent que les professeurs considèrent leurs collègues comme une source majeure d'idées, de conseils et d'expertise en évaluation. À l'instar de ces chercheurs, nous déplorons que trop d'enseignants ne prennent pas le temps de solliciter et d'encourager la collaboration de leurs collègues lorsque vient le temps de planifier et d'effectuer l'évaluation des apprentissages.

Dans les départements, de nombreux enseignants sont compétents en évaluation et ces personnes devraient être identifiées et mises à contribution comme ressources pour leurs collègues. Sur ce point, nous savons que de nombreux professeurs sont disposés à aider leurs collègues en leur faisant partager leur expertise et que plusieurs autres auraient avantage à profiter des compétences de ces collaborateurs. Mais nous savons aussi que la peur d'être jugés comme incompétents constitue un obstacle majeur à la collaboration entre pairs, nous renvoyant encore à la problématique de l'absence de formation en pédagogie. Les lacunes en formation pédagogique préalable et les dynamiques interpersonnelles dans les départements peuvent probablement favoriser l'existence de la peur d'être jugés, ce qui peut décourager certains professeurs de demander conseil ou de faire valider, par exemple, la clarté et la pertinence de leurs questions d'examens ou leurs consignes de travaux longs.

Recommandation

Recommandation 23: Avant de jumeler un professeur nouvellement embauché avec un professeur en exercice, nous recommandons de faire l'évaluation des pratiques et des croyances du professeur en exercice, afin de favoriser que le nouveau professeur soit guidé adéquatement vers des pratiques correctes.

La nature de la formation

Pour une politique de la formation des maîtres

Toutes les études qui ont été faites depuis une dizaine d'années, tant au Québec qu'aux États-Unis et en Europe, concourent à démontrer que l'enseignement est une profession extrêmement complexe et que le champ de l'évaluation des apprentissages, à lui seul, constitue une bonne part de cette complexité. La tâche est lourde, difficile et souvent ingrate. Les enseignants ont besoin d'une formation spécifique à l'exercice de leur profession et nous pensons que les enseignants, qui se donnent une formation en pédagogie, contribuent à mettre leur profession en valeur.

Comme l'enseignement, l'évaluation des apprentissages n'est pas seulement un art. Il faut posséder des compétences techniques et des connaissances conceptuelles qui appuient le jugement, qui facilitent la résolution de problèmes et qui contribuent à développer un sentiment de sûreté, de compétence. Pour ne pas corrompre le processus de l'évaluation et ainsi compromettre la validité et l'interprétabilité des résultats de l'évaluation, les professeurs doivent savoir éviter les nombreux biais qui mènent parfois à l'indulgence ou à l'intransigeance injustifiées.

À cet égard, nous rappelons ici que le Conseil supérieur de l'éducation, dans son "Rapport annuel 1990-1991 sur l'état et les besoins de l'éducation", recommandait que des politiques de la formation des maîtres soient mises en oeuvre:

> "Le Conseil croit que tant qu'une politique (de la profession enseignante) ne sera pas adoptée, la situation de la formation de maîtres continuera d'évoluer de façon anarchique, (...) auquel cas la société québécoise de demain tout entière sera gravement pénalisée" (p. 46).

Dans l'esprit de cette recommandation, le Conseil décrivait quelques balises qui peuvent guider une politique gouvernementale ou institutionnelle de la formation des maîtres:

- l'exigence, dès l'embauche, d'une formation didactique et pédagogique concomitante de la pratique professionnelle, au cours des premières années d'enseignement (ce pourrait être, par exemple, au cours des cinq premières années d'enseignement, années pendant lesquelles s'acquiert l'identité professionnelle);
- le développement de formules de perfectionnement pédagogique sur le modèle de l'expérience Performa;
- l'intégration d'une forme ou l'autre de perfectionnement à la tâche des enseignants du collégial (C.S.E., 1991, p. 46).

Recommandation

Recommandation 24: Nous nous associons au Conseil supérieur de l'éducation en recommandant que les collèges se donnent une politique institutionnelle de la formation des maîtres qui puisse favoriser le développement et le maintien des compétences minimales en mesure et évaluation des apprentissages, chez tous les professeurs de cégeps, anciens comme nouveaux.

Recommandations

Recommandation 25: Nous recommandons qu'une formation minimale en pédagogie et spécifiquement en mesure et évaluation soit rendue obligatoire pour tous les nouveaux professeurs embauchés au collégial et que l'atteinte de cette formation minimale fasse l'objet d'échéances prescrites.

Dans l'un des groupes réacteurs, nous avons testé l'hypothèse de recommander qu'une formation minimale soit rendue obligatoire pour les professeurs en exercice. Si plusieurs disent souhaitable d'obliger les professeurs actuels à suivre une formation en évaluation des apprentissages, les autres reconnaissent que toute obligation donnerait lieu à de fortes oppositions. "Les profs sont méfiants. Si on les oblige, ils vont se braquer". Selon plusieurs participants, la révision des programmes ou la mise à jour des politiques des établissements sur l'évaluation des apprentissages peuvent susciter l'intérêt des professeurs à l'égard des principes de mesure et évaluation.

Recommandation 26: Nous recommandons que tous les professeurs en exercice soient activement invités à développer et à maintenir des compétences minimales en mesure et évaluation et à participer à des sessions de perfectionnement touchant les six champs de compétence décrits ici.

La tâche de formation, si complexe soit-elle, n'est pas insurmontable. Afin d'organiser cette formation ou pour mettre sur pied des services destinés aux enseignants et aux autres personnels, il faut notamment:

1. Identifier les personnes-ressources qui font déjà de la formation en mesure et évaluation ou qui peuvent contribuer à la formation des personnels;

2. Mettre sur pied une équipe de formateurs;

3. Mettre au point un programme de formation pour les professeurs et pour les administrateurs;

4. Établir les budgets et obtenir l'aide financière appropriée;

5. Diversifier les modes de formation en faisant appel, notamment, à l'expertise disponible en matière d'andragogie, de formation à distance, de technologie éducative: exploitation de cassettes vidéo ou audio, télévision éducative, modules d'autoformation, etc.

6. Offrir de la formation, tant dans le cadre formel de cours universitaires, que dans des cadres plus informels comme des journées pédagogiques dédiées, des séminaires départementaux ou inter-départementaux, des bulletins pédagogiques d'établissement.

Nonobstant la problématique évidente des budgets de formation, il faut rappeler qu'il y a actuellement douze mille professeurs en exercice au collégial. Cela pose le problème de la disponibilité, au Québec, d'un nombre suffisant de formateurs qualifiés en mesure et évaluation.

Un cadre favorable au développement des compétences: L'évaluation de l'enseignement

Plusieurs recherches ont démontré que beaucoup de professeurs croient fermement que leur rôle est de transmettre de l'information et du savoir. Ces recherches ont aussi démontré que beaucoup d'enseignants pensent, en début de carrière, qu'ils seront de meilleurs professeurs que ceux qui les ont précédés (Pajares, 1992, p. 323).

Ces croyances de débutants sont probablement présentes au collégial. Or, le réseau des collèges ne donne pas de feedback à ses professeurs. Ceux-ci peuvent enseigner pendant des années sans recevoir de commentaires évaluatifs, ni négatifs ni positifs, sur leur enseignement. Cette absence de rétroaction peut très facilement faire croire à certains que leur

enseignement est considéré comme adéquat, ce qui ne peut que renforcer leur croyance qu'ils sont compétents. Même si la plupart de nos professeurs n'ont pas de formation en mesure et évaluation, on a vu que la grande majorité se jugent compétents (question 56 du questionnaire) et cela ne contribuera, probablement pas, à ce que beaucoup d'entre eux soient disposés à se remettre en question et à écouter un discours, les incitant à se mettre à jour en mesure et évaluation des apprentissages.

S'ils se croient compétents, ceux qui ne le sont pas vont attribuer à d'autres causes leurs difficultés en évaluation. Par exemple, certains attribueront la difficulté de certaines tâches, comme la correction des travaux d'écriture, au manque de temps. De là, ils demanderont qu'on allège leur tâche mais ils n'auront pas nécessairement tendance à ré-examiner leurs pratiques.

Dans un groupe réacteur, un enseignant nous a affirmé que ce sont les étudiants qui vont pousser les professeurs à se perfectionner et cela se produira lorsque les étudiants feront l'évaluation de l'enseignement.

Recommandation

Recommandation 27: L'absence de rétroaction sur la pratique enseignante peut être trompeuse et peut, à la limite, permettre à de fausses croyances de se cristalliser. C'est pourquoi nous recommandons que les collèges donnent à leurs enseignants l'occasion de recevoir des commentaires constructifs sur leurs pratiques pédagogiques, et ce, dans le cadre d'un processus structuré d'évaluation de l'enseignement.

Mettre au point un curriculum de formation en mesure et évaluation

Plusieurs recherches, évoquées au chapitre 2, font état de l'insatisfaction des professeurs à l'endroit des contenus des cours universitaires qui leur sont offerts en mesure et évaluation. De plus, dans les groupes réacteurs, des professeurs affirment qu'à cause de la très grande diversité de croyances et de pratiques, on ne peut pas proposer un seul modèle de formation, un seul curriculum. Si la formation est adaptée, disent-ils, les gens vont la demander.

Dans son rapport de 1993, le comité conjoint de la Fédération des cégeps sur le perfectionnement décrit cinq étapes de réalisation de tout processus de développement des compétences. À ce point de notre recherche, nous avons réalisé les deux premières étapes, à savoir:

1. Identifier les problèmes et les besoins;

2. Analyser les besoins de développement des compétences;

Selon le cadre de référence proposé par ce comité conjoint (Fédération des cégeps, 1993, p. 13), les trois étapes suivantes seraient:

3. Concevoir des activités de développement des compétences;

4. Mettre en oeuvre un plan d'activités de développement des compétences;

5. Évaluer et réinvestir les apprentissages.

Recommandations

Recommandation 28: Nous recommandons qu'on mette au point un curriculum de formation en mesure et évaluation et que ce curriculum soit adapté aux besoins de l'enseignement collégial, suffisamment souple pour convenir aux divers profils de formation et d'expérience des professeurs de cégeps et ajusté aux rôles que nous entendons confier à l'évaluation dans les écoles d'aujourd'hui. Ce curriculum de formation devra définir des critères qui permettront d'évaluer les activités qui auront été mises en oeuvre dans le cadre du développement des compétences en mesure et évaluation des apprentissages.

Recommandation 29: Dans la mesure où nous souhaitons que les professeurs de cégeps fréquentent des cours universitaires sur le sujet de l'évaluation des apprentissages, nous recommandons que tous les formateurs d'enseignants s'assurent que la formation, qu'ils offrent aux professeurs des collèges, soit pertinente et adaptée à la réalité de l'enseignement collégial.

À la lumière des résultats de la présente recherche, la mise au point d'un curriculum de formation en mesure et évaluation devrait tenir compte, notamment, des profils de réponses par strates, en annexe 5, afin d'y cibler des strates qui auraient des besoins prioritaires dans certains domaines.

6.3.4 Cadres et conseillers pédagogiques: Perfectionnement en évaluation des apprentissages

Dans les collèges, les cadres des services éducatifs et les conseillers pédagogiques jouent des rôles majeurs au chapitre des pratiques en évaluation des apprentissages. Il suffit de lire les politiques institutionnelles d'évaluation des apprentissages pour y constater qu'ils ont des responsabilités spécifiques à cet égard. Ils peuvent influencer la création, le maintien ou

le changement de certaines croyances ou de certaines pratiques. Des choses peuvent changer, selon qu'ils exercent ou non leur leadership. Ils disposent de moyens pour changer des choses mais ils peuvent aussi favoriser des statu-quo.

Si les professeurs se donnent de la formation ou du perfectionnement en mesure et évaluation, il faut que les administrateurs et les conseillers pédagogiques développent parallèlement leurs propres compétences, pour comprendre ce que font les professeurs et pour les aider à faire ce qu'ils ont à faire. Rappelons d'ailleurs que les réponses à la question 58 de notre questionnaire révèlent que plusieurs répondants souhaitent que des ressources humaines soient mises à contribution pour les aider, individuellement ou en groupe, à parfaire leur compétence en évaluation ou à résoudre des problèmes ponctuels. Les conseillers pédagogiques, notamment, devraient pouvoir répondre à des demandes de counselling individuel en mesure et évaluation. Les cadres et les conseillers pédagogiques devraient posséder une compréhension minimale des concepts fondamentaux en mesure et évaluation, afin d'éviter toute dissonance dans le discours des uns et des autres.

Étant fréquemment d'anciens professeurs de cégeps, les cadres et les conseillers pédagogiques n'ont pas nécessairement eu, non plus, une formation universitaire minimale en pédagogie et en mesure et évaluation. De plus, leurs responsabilités professionnelles ne leur permettent probablement pas de se tenir à jour au chapitre des concepts et des pratiques en évaluation des apprentissages.

Voici un exemple extrait d'une rencontre avec l'un des groupes réacteurs. Lors d'une discussion sur la pratique d'attribuer des points pour l'effort consenti ou pour la participation en classe, un administrateur de collège se joignit aux enseignants qui affirmaient que ces pratiques sont, en quelque sorte, incontournables; "Quand j'étais prof, je le faisais moi aussi" ou "La théorie, c'est bien beau, mais en salle de classe, c'est une autre histoire". Par ce genre de réponses, cet administrateur donnait une apparence de normalité à des pratiques fort discutables. Etait-ce par

complaisance? Etait-ce par peur de désavouer les pratiques qu'il avait lui-même comme professeur?

À notre avis, lorsqu'un administrateur banalise ouvertement une pratique discutable, en disant qu'il faisait ça aussi "dans le temps" ou que les professeurs n'ont pas le choix de le faire (par exemple, pour motiver les étudiants), il fait deux choses du même coup. D'abord, il dit implicitement qu'il n'entend pas investir de l'énergie pour participer à des changements là-dessus, jugeant que ces pratiques sont acceptables et, du même coup, que des changements sont non nécessaires. Ensuite, il participe à confirmer, auprès des professeurs, que ces pratiques sont acceptables ou à tout le moins tolérables et ce, à cause de circonstances hors du contrôle du professeur en classe. Au fond, il y a ici un double discours. Devant le chercheur, on admet que telle pratique n'est pas correcte au vu des arguments théoriques. Mais on ajoute tout de suite qu'au fond, les professeurs n'ont pas beaucoup le choix et que cette pratique est "malheureusement" normale.

Recommandation

Recommandation 30: Les cadres des services éducatifs des collèges et les conseillers pédagogiques devraient être invités à se donner du perfectionnement ou à maintenir des compétences minimales en mesure et évaluation, afin qu'ils puissent être en mesure de collaborer utilement à l'application des politiques locales d'évaluation des apprentissages, ainsi qu'aider les professeurs à résoudre des problèmes de mesure ou d'évaluation.

6.3.5 Les collaborateurs à ces formations

Les autres grands partenaires de la formation de la clientèle collégiale, notamment la Direction générale de l'enseignement collégial, la Fédération des cégeps ainsi que les syndicats, doivent aussi poser des

gestes visibles et concrets qui favorisent le développement des compétences en mesure et évaluation et qui contribuent à maintenir la crédibilité des pratiques des professeurs. Nous croyons que les syndicats de professeurs ont un pouvoir et une responsabilité de premier plan au chapitre de la protection de la crédibilité des pratiques des professeurs et du maintien de la confiance du public. De plus, nous nous associons au comité conjoint de la Fédération des cégeps sur le perfectionnement lorsqu'il affirme le principe fondamental suivant:

> "... il est de la responsabilité des employés des collèges de se préoccuper du développement de leurs compétences, d'y consentir les efforts nécessaires et de faire en sorte que les utilisateurs des services qu'ils donnent profitent des retombées des activités de développement des compétences" (Fédération des cégeps, 1993, p. 13).

En conséquence, nous pensons que les syndicats de professeurs ainsi que les autorités des collèges doivent contribuer à diffuser et à appuyer explicitement et activement ce principe fondamental.

Recommandation

Recommandation 31: Les syndicats de professeurs devraient contribuer activement à protéger la crédibilité des pratiques des professeurs en matière d'évaluation et de notation et ce, en encourageant ces derniers à respecter des seuils élevés de compétence et à poser des gestes visibles concourant à maintenir la confiance du public.

6.3.6 Services à offrir

La question 58 de notre questionnaire a investigué les ressources et services qui pourraient être utiles aux enseignants des cégeps. On a vu, au chapitre 5, que les professeurs souhaitent, en toute priorité, qu'on mette à leur disposition des modèles qui pourront les aider à résoudre des

problèmes pratiques en évaluation: modèles d'examens et modèles de consignes de travaux écrits. Ensuite, les répondants souhaitent de l'aide informatisée: des banques d'items, des systèmes de correction et d'analyse des examens. Enfin, les professeurs désirent avoir accès à des ressources humaines, soit pour alléger le fardeau des corrections, soit pour leur donner du counseling en matière d'évaluation des apprentissages. Un professeur d'un groupe réacteur est surpris du palmarès obtenu à la question 58 parce qu'en principe, les professeurs "moyens", dans le réseau collégial, sont plus expérimentés et plus âgés. C'est pourquoi ce professeur s'étonne que les répondants demandent des modèles de ci et de ça.

Plusieurs enseignants agissent en solitaires lorsqu'ils préparent leurs dispositifs d'évaluation. Ils consultent très rarement leurs collègues pour valider, a priori, leurs examens ou autres consignes servant à évaluer les apprentissages. Or, plusieurs expriment le souhait d'obtenir des modèles (des recettes?) qui viendraient de l'extérieur. On peut s'inquiéter à double titre. D'une part, la disponibilité de modèles pourrait ne pas favoriser les échanges entre professeurs, en permettant à certains de ne pas déroger à leur isolement. D'autre part, on pourra craindre que des enseignants ne puisent des modèles que ce qui correspond à leurs pratiques ou à leurs croyances, puisque nous savons que les croyances ont tendance à s'autoperpétuer.

Recommandation

Recommandation 32: Nous recommandons que plusieurs services et ressources soient offerts aux professeurs des cégeps dont, notamment, les ressources demandées à notre question 58. Cependant, la mise au point de modèles doit s'accompagner de stratégies favorisant la concertation entre professeurs. De plus, ces éventuels modèles d'examens ou modèles de consignes devraient être assez diversifiés pour favoriser, chez certains professeurs, une remise en question des croyances qui nécessitent des changements.

CONCLUSION

Selon Chilcott (1992, p. 31), une bonne conclusion de recherche consiste à répondre à la question: "et alors?" ("so what!"). La recommandation de cet éminent anthropologue nous convient bien et nous proposons une conclusion en trois temps, dans les axes suivants: ce qu'on voulait faire, ce qu'on a fait et ce qui reste à faire.

1. Ce qu'on voulait faire

Cette recherche est fondée sur la question des compétences des professeurs des cégeps en mesure et évaluation des apprentissages en salle de classe. Les enseignants des cégeps ont l'entière responsabilité de l'évaluation des apprentissages de leurs étudiants. Mais pour accomplir cette tâche, ils disposent de formations inégales, sinon aléatoires, et de nombreux professeurs n'ont aucune formation pouvant les guider lors de l'évaluation des apprentissages de leurs étudiants. À l'embauche, on ne requiert pas qu'ils aient une formation en pédagogie.

Nous savons que l'évaluation des apprentissages en salle de classe peut aider à enseigner et à apprendre. Nous savons aussi que l'évaluation des apprentissages transporte le pouvoir inhérent de déterminer des conséquences éducatives et sociales dont les répercussions se font sentir, hors de l'école, sur de longues périodes. Nous savons, enfin, que l'évaluation occupe une portion importante de la tâche des enseignants ainsi que du temps de classe et constitue pour plusieurs un fardeau fastidieux.

Tel que le démontrent nos chapitres 2 et 3, les impacts de l'évaluation des apprentissages sont si importants qu'on peut en arriver à modifier les pratiques d'enseignement simplement en changeant les pratiques d'évaluation.

À partir de ces constats, nous cherchons à définir le domaine de la compétence en mesure et évaluation, du point de vue des professeurs des cégeps, afin de guider la mise au point d'activités de perfectionnement à l'intention des professeurs. Puisque les croyances sont en amont des

pratiques et que toute démarche visant à modifier les pratiques doit tenir compte des croyances, tout projet d'intervention sur les pratiques d'évaluation des apprentissages doit être fondé sur une bonne connaissance des croyances des professeurs à l'égard de l'évaluation.

2. Ce qu'on a fait

Tel que prévu dès le départ, cette recherche fait la description des croyances et des pratiques des professeurs en matière de mesure et évaluation des apprentissages en salle de classe.

> "Si un champ d'étude ne possède ni limites ni paramètres clairement définis, on commence en principe par décrire cet univers. (...) On ne sait pas quelles variables sont particulièrement significatives dans les milieux naturels tant que nous ne disposons pas de descriptions détaillées de leurs caractéristiques et de leur influence. En l'absence de telles données descriptives, on choisit d'étudier les mauvais problèmes, on teste des hypothèses non pertinentes et on établit des inférences erronées" (Miles et Huberman, 1991, p. 447).

C'est dans cet esprit que nous avons maintenu et atteint nos trois objectifs de départ, faisant appel, notamment, à une méthodologie qualitative pour valider les données quantitatives recueillies. Cette recherche permet d'organiser le domaine d'étude de l'évaluation des apprentissages au collégial, d'abord sous deux grands paramètres: les croyances et les pratiques, et ensuite, dans chacun de ces paramètres, sous une analyse en six grandes catégories, inspirées des catégories de Stiggins, en annexe 7. Il était nécessaire de proposer une telle structure d'étude parce que le domaine de l'évaluation des apprentissages est peu étudié au collégial et que nous voulons guider la mise au point d'activités de perfectionnement dans ce domaine.

Le partage de l'information

Notre recherche génère une masse considérable d'information. Or, nous savons que partager l'information, c'est partager le pouvoir de l'information (Guba et Lincoln, 1991, p. 369). C'est pourquoi un tel rapport de recherche peut à lui seul disloquer l'équilibre et créer de l'anxiété. Nos objectifs étaient et restent clairs: cette étude descriptive concourt à mettre en lumière des pistes d'intervention qui devraient aider les professeurs à développer et maintenir leurs compétences en évaluation des apprentissages. Le développement de leurs compétences permettra aux professeurs de mieux assumer la tâche de l'évaluation en salle de classe ainsi que conserver et protéger la pleine autonomie dans laquelle ils exercent cette responsabilité. Dans cet esprit, notre recherche sera utile tant aux conseillers pédagogiques des cégeps qu'aux professeurs eux-mêmes:

a. Les conseillers pédagogiques seront maintenant mieux informés sur les profils de croyances et de pratiques, ce qui les guidera dans leurs interventions de formation et de perfectionnement. Ils comprendront surtout que les professeurs ont besoin d'une aide pratique, adaptée à leurs besoins, plutôt que de cours théoriques en psychométrie ou en statiques. Avec une attitude de collaboration plutôt qu'une approche dogmatique, les conseillers pédagogiques peuvent guider une réflexion sur les croyances et les pratiques en évaluation de façon à ce que les pratiques soient conformes aux principes admis en pédagogie et en évaluation, tout en étant adaptées à la réalité de l'ordre collégial.

b. Les professeurs auront décrit, ici, des pratiques et croyances en évaluation, un peu comme dans un processus de reddition de compte, avec transparence. Ils restent responsables de leurs pratiques mais nous avons plus d'information qu'avant sur la conduite de ces pratiques. Il leur appartient toujours de pouvoir gérer les connaissances et les croyances sur lesquelles ils fondent leurs pratiques mais ils peuvent prendre l'initiative de questionner certaines croyances

pour établir dans quelle mesure elles correspondent ou non aux connaissances empiriques.

Les recommandations

Les tableaux descriptifs de l'ensemble de l'échantillon, les tableaux comparatifs par strates, les analyses factorielles et les commentaires extraits des groupes réacteurs permettent d'identifier des croyances et des pratiques qui illustrent certaines lacunes dont, tout particulièrement:

- la méconnaissance de l'évaluation formative;
- l'utilisation de l'évaluation à des fins détournées ainsi que la pollution des notes versées au bulletin;
- le caractère facultatif de la formation des professeurs en mesure et évaluation.

Le rapport de recherche propose plusieurs recommandations portant sur la correction de ces lacunes, sur le développement des compétences des professeurs ainsi que sur des services à offrir.

3. Ce qui reste à faire

L'utilisation de l'information

La diffusion

Nous souhaitons que nos résultats de recherche fassent l'objet de la plus large diffusion possible auprès des professeurs des cégeps. Dans chaque catégorie de compétence, ceux-ci devraient étudier leurs propres croyances et pratiques à la lumière de celles que nous décrivons ici. Les professeurs pourraient, notamment, étudier nos tableaux d'analyses stratifiées pour en tirer les informations concernant leur strate d'appartenance. Nous souhaitons que les enseignants se servent de ces tableaux

pour guider et enrichir une réflexion pédagogique dans leurs départements, par exemple, dans des journées pédagogiques dédiées.

Nos résultats de recherche devraient être aussi diffusés auprès des conseillers pédagogiques et de la direction des études des collèges qui pourraient s'en servir (et c'est là notre objectif fondamental) pour planifier des activités contribuant au développement des compétences en mesure et évaluation dans les établissements.

La mise en oeuvre des recommandations

Nos analyses, au chapitre 6, donnent lieu à plusieurs recommandations qui intéressent, en toute priorité, les enseignants des cégeps qui s'engagent dans un processus d'analyse réflexive de leur tâche et de leurs besoins de perfectionnement. La mise en oeuvre de certaines de nos recommandations relève des responsables de la formation et du perfectionnement. Elles devraient guider ceux-ci dans la mise au point d'activités de formation à l'intention des enseignants des collèges. Nous pensons ici, tout particulièrement, aux répondants locaux de PERFORMA qui, dans leurs collèges respectifs ou dans leur assemblée générale de réseau, pourraient profiter des structures que nous proposons ici ainsi que des données de recherche pour établir une programmation d'activités à offrir aux professeurs.

D'autres recommandations interpellent les directions des collèges ou les syndicats des professeurs car leur mise en oeuvre exige des choix politiques ou administratifs qu'ils ont les moyens de faire.

Pour passer à l'action, les gens qui voudront adopter nos recommandations devront établir des stratégies qui favoriseront le changement. Ces stratégies devront tenir compte de la nécessité incontournable d'agir sur les croyances avant d'espérer agir sur les pratiques. Dans cet esprit, Shavelson et Stern (1981) suggèrent une chaîne d'événements qui devrait être mise en oeuvre pour que les professeurs adoptent des résultats de recherche:

- les professeurs doivent prendre conscience des diverses croyances personnelles (subjective beliefs) qu'ils entretiennent sur l'enseignement et sur l'évaluation;
- ces croyances doivent être confrontées à des vérifications empiriques. Les professeurs doivent pouvoir valider leurs croyances, soit en prenant connaissance de résultats de recherches pratiques, soit en effectuant eux-mêmes des expérimentations en salle de classe (Cross et Angelo, 1988);
- des croyances personnelles se transforment en savoirs objectifs si on peut les confirmer empiriquement;
- des croyances personnelles peuvent être abandonnées si elles sont infirmées par des démonstrations empiriques.

Un tel scénario mérite d'être testé, tenant compte de la nature des croyances (chapitre 2) et de leur caractère persistant.

Maintien de la responsabilité enseignante

Les professeurs et les départements ont toujours été responsables de la pratique de l'évaluation des apprentissages des étudiants et les politiques institutionnelles d'évaluation des apprentissages ont confirmé cette responsabilité. Nous croyons que les professeurs doivent conserver la responsabilité de la pratique de l'évaluation des apprentissages en salle de classe car nous croyons que leur pleine autonomie, dans ce geste clé du processus de sanction des études au collégial, est pédagogiquement justifiée. D'ailleurs, les choix qui sont faits dans le cadre des mesures de renouveau de l'enseignement collégial excluent la mise sur pied d'un système d'examens nationaux dont on rappelle qu'on s'est naguère affranchi avec conviction (Ministère de l'Enseignement supérieur et de la Science, avril 1993, p. 27). Cependant, nous croyons que le maintien et la protection de cette pleine autonomie passe par une démonstration visible

et crédible de la compétence professionnelle des professeurs et que cette compétence, loin d'être définitivement acquise et statique, doit faire l'objet de mises à jour pertinentes et régulières. C'est dans l'esprit de cette recherche de crédibilité publique que la ministre de l'Enseignement supérieur et de la Science a fait de l'évaluation sa priorité de travail en 1992-1993 et a articulé une réforme majeure de l'enseignement collégial, dont la création d'une commission d'évaluation de l'enseignement collégial s'avère une pièce maîtresse.

L'utilisation de notre questionnaire

Nous pensons que notre questionnaire de recherche, en annexe 1, pourrait être mis à contribution pour diagnostiquer des besoins de formation des professeurs. Par exemple, il pourrait être administré à plusieurs professeurs qui enseignent un même cours, si ceux-ci désirent comparer leurs pratiques d'évaluation des apprentissages ou leurs conceptions de l'enseignement et de l'évaluation. On pourrait aussi s'en servir à profit dans le cadre d'une activité de perfectionnement, par exemple au CPEC-PERFORMA, pour comparer les croyances et les pratiques d'un groupe de professeurs aux résultats nationaux publiés dans la présente recherche. Enfin, notre questionnaire pourrait servir dans d'autres recherches traitant, par exemple, des conceptions de l'enseignement ou de l'évaluation.

Des recherches futures

Dans cette recherche, nous avons décrit l'univers de ce champ d'étude que constitue l'évaluation des apprentissages en salle de classe. On dira peut-être que, dans les limites de ce champ d'étude, nous avons fait une écologie de la tâche des professeurs. Mais plusieurs questions restent ouvertes et, suite à notre étude, plusieurs autres pistes de recherche devraient être explorées. Par exemple, il nous reste à savoir:

- quelles sont les croyances des étudiants à l'égard de l'évaluation des apprentissages et comment ces croyances font pression sur les pratiques des professeurs;
- quelles sont les opinions des étudiants à l'égard des pratiques des professeurs;
- quelles sont les croyances des cadres et des conseillers pédagogiques, dans les services éducatifs, à l'égard de l'évaluation des apprentissages et comment ces croyances interviennent dans les pratiques d'encadrement et dans les services offerts aux professeurs.

De plus, dans la présente recherche, le problème de l'évaluation en laboratoire et en milieu de stage est exclu, car ces milieux imposent des contraintes différentes de celles de la salle de classe. À cet égard, nous pensons que la problématique de l'évaluation en laboratoire et en milieu de stage devrait faire l'objet d'une étude spécifique.

Le processus du développement des compétences

On sait que les croyances à l'égard de l'évaluation sont en amont des pratiques d'évaluation. On sait aussi que les pratiques d'évaluation ont des effets, désirables et non désirables, tant sur l'enseignement (la tâche du professeur, la planification, la qualité de l'enseignement) que sur l'apprentissage (l'étude, la qualité de l'apprentissage, la persistance cognitive). Si notre étude permet de mieux connaître cet aspect de l'enseignement collégial et donne quelques pistes de perfectionnement, d'autres études sont requises pour définir les conditions qui doivent être mises en oeuvre pour que les efforts de perfectionnement donnent véritablement lieu à des changements utiles et satisfaisants.

Les croyances, les relations intra-départementales, la "culture locale" des cégeps, les pressions des étudiants, les enjeux politiques de la liberté

académique sont autant de facteurs qui compliquent le problème de la prestation d'activités de perfectionnement. Ces facteurs s'ajoutent aux dimensions cognitives du développement des compétences, dont, notamment, les conceptions de l'enseignement et les conceptions de l'évaluation. Il est, désormais, naïf de penser qu'on peut tout simplement "organiser des cours pour les profs". Afin de mettre en oeuvre des scénarios efficaces de développement des compétences, il faut connaître et comprendre comment les professeurs ont appris ce qu'ils savent, comment ils ont adopté des croyances à l'égard de l'évaluation des apprentissages et comment ils perçoivent leurs pratiques. À cet effet, une recherche de nature qualitative devrait être mise en oeuvre pour décrire les facteurs qui favorisent ou défavorisent la formation en cours d'emploi, le perfectionnement, le changement de croyances, le changement de pratiques. Cette recherche devrait tout particulièrement distinguer les sources d'influence sur les croyances, selon qu'elles sont endogènes ou exogènes.

La vie départementale

Nous croyons pertinent de mieux connaître les contributions de la vie départementale à l'évolution ou à l'absence d'évolution des croyances et des pratiques des professeurs. On sait, en effet, que le département est un lieu de transmission des savoir-faire en mesure et évaluation. Mais il peut être aussi un lieu de tensions, favorisant des attitudes de retrait ou de statu-quo et inhibant le processus de changement.

Les fondements des pratiques

Il nous faut dépasser le stade de la recherche descriptive et amorcer des études évaluatives afin d'identifier clairement les croyances et les pratiques qui ont des influences défavorables sur:

- la crédibilité des notes et du processus de sanction des études;
- la tâche des enseignants;
- la qualité de l'enseignement et de l'apprentissage.

Ces recherches évaluatives feraient émerger, peut-être par le processus qualitatif de l'entrevue ou du groupe de tâche, des pratiques pédagogiques qui pourraient être qualifiées de bonnes ou de mauvaises. Un tel processus permettrait d'investiguer plus à fond les motifs sous-jacents des professeurs à l'égard de certaines pratiques, par exemple: la pratique de donner des points pour l'exécution de devoirs "à la maison", la pratique d'accorder la même note à tous les participants d'un travail en équipe, la pratique de donner une note de passage à un étudiant faible lorsque ses efforts semblent appréciables, l'isolement des professeurs dans la confection de leurs dispositifs d'évaluation.

La question des croyances et des attitudes

Cette recherche décrit de nombreuses croyances dont on sait qu'elles sont en amont des pratiques et sur lesquelles il faut intervenir si on cherche à agir efficacement sur les pratiques. À partir des types de croyances, décrits au chapitre 2, une autre recherche devrait être mise en oeuvre pour classer les croyances des professeurs selon une typologie pertinente et, de là, définir des stratégies d'intervention différenciées, en fonction de chaque type de croyance.

Par ailleurs, nous savons que les conceptions de l'enseignement jouent certainement des rôles dans la genèse des croyances. Dans cette perspective, il faut étudier et mieux comprendre les conceptions que les professeurs se font, tant de l'enseignement que de l'évaluation, afin de déterminer quels types de relations existent entre ces deux objets et comment ces conceptions jouent sur la genèse des croyances.

Il y aurait aussi lieu d'explorer la question de la formation des attitudes et d'y tester notre modèle, très simple, des trois temps de la prise de décision (voir ce modèle au chapitre 2) pour voir en quoi il peut contribuer à expliquer la formation des attitudes.

Ces pistes de recherche sont loin d'être exhaustives mais elles ont toutes trait au processus de la formation professionnelle des enseignants des cégeps ainsi qu'à la nécessaire qualité de l'enseignement. Tout comme la recherche que nous terminons ici, ces futures recherches peuvent contribuer à conférer à la profession enseignante toute la crédibilité et le respect qu'elle mérite.

BIBLIOGRAPHIE

Agnew, E. J. (1985). The grading policies and practices of high school teachers. [Communication présentée dans le cadre du colloque de L'AERA à Chicago] (50 p.).

Airasian, P. W. (1991). Perspectives on measurement instruction. Educational Measurement: Issues and Practice, 10(1), 13-16.

Alverno College (1985). Assessment at Alverno College. Milwaukee: Alverno College.

American Federation of Teachers (1990). Standards for teacher competence in educational assessment of students. Educational Measurement: Issues and Practice, 9(4), 30-32.

Anderson, J. O. (1989, juin). Evaluation of student achievement teacher practices and educational measurement. The Alberta Journal of Educational Research, 35(2), 123-133.

Applebee, A. N. & Langer, J. A. (1988). Speaking of knowing: conceptions of learning in academic subjects. Academic learning in high school subjects. Final report. Office of educational research and improvement. Washington.

Barnes, S. (1985). A study of classroom pupil evaluation: The missing link in teacher education. Journal of teacher education, 36(4), 46-49.

Barre-de-Miniac, C. (1985). Professeurs, élèves, parents face à l'évaluation. Paris: Institut national de recherche pédagogique (Collection Rapports de recherches, no. 4).

Bigard, A., & Guyot, Y. (1980). Attitudes à l'égard des examens dans l'enseignement supérieur. Revue internationale de psychologie appliquée, 29, 355-372.

Bloom, B. S. et al (1979). Caractéristiques individuelles et apprentissages scolaires. Paris: Nathan.

Boothroyd, R., McMorris, R. & Pruzek, R. (1992, avril). What do teachers know about testing and how did they find out? [Communication présentée dans le cadre du colloque de L'AERA à San-Francisco].

Brookhart, S. M. (1991). Grading practices and validity. Educational Measurement: Issues and Practice, 10(1), 35-36.

Burelle, R., Gadbois, L., Parent, C. & Séguin, S. (1991). Mesurer des croyances et des attitudes pour mieux comprendre des comportements. Mesure et évaluation en éducation, 14(1), 61-77.

Campbell, M. M. & Felx, C. L. (1988). Auto-évaluation: concept et pratiques. Montréal: Université de Montréal, Services pédagogiques.

Cardinet, J. (1986). Evaluation Scolaire et Pratique. Bruxelles: De Boeck Université.

Chase, C. I. & Wakefield, L. M. (1984). Testing and grading: Faculty practices and opinions. Indiana Studies in Higher Education Monograph No. 51, Indiana Univ., Bloomington (18 p.).

Chilcott, J. H. (1992, mars). Some contributions of anthropology to professional education. Educational Researcher, 21(2), 31-35.

Commission de l'évaluation, Conseil des Collèges (1990, mars). La mise en oeuvre des politiques institutionnelles d'évaluation des apprentissages: Rapport 1988-89. Québec: Gouvernement du Québec.

Conseil des Collèges, R. Grégoire et al. (1985) Étude de la pratique professionnelle des enseignants et enseignantes de cégep. Québec: Gouvernement du Québec.

Conseil des Collèges (1989). L'harmonisation du secondaire et du collégial: l'état et les besoins de l'enseignement collégial. Québec: Gouvernement du Québec.

Conseil supérieur de l'éducation (1991). La profession enseignante: vers un renouvellement du contrat social. Québec: Gouvernement du Québec.

Crooks, J. J. (1988). The impact of classroom evaluation practices on students. Review of Educational Research, 58, 438-481.

Cross, L., Frary, R. & Weber, L. (1992). College grading : Achevement, attitudes and effort. Blacksburg, Virginia: Virginia Polytechnic Institute and State University.

Cross, P. & Angelo, T. A. (1988). Classroom assessment techniques: a handbook for faculty.: University of Michigan. NCRIPTAL.

Crowl, T. K. & Berkowitz, M. (1985). Regular and special education teachers'marking attitudes and behavior. Contemporary Educational Psychology, 10, 260-267.

Crowl, T. K. (1984). Grading behavior and teachers' need for social approval. Education, 104(3), 291-295.

Daunais, J. (1986). L'entretien non directif. Benoit Gauthier (directeur), Recherche Sociale, (chapitre 11). Québec: PUQ.

DeKetele, J. & Roegiers, X. (1991). Méthodologie du recueil d'informations. Bruxelles: De Boeck Université.

DeLandsheere, V. (1988). Faire réussir, faire échouer: La compétence minimale et son évaluation. Paris: P.U.F.

Doyon, C. & Juneau, R. (1991). Faire participer l'élève à l'évaluation de ses apprentissages. Laval: Beauchemin (Collection Agora).

Ducharme, S. & Forcier, P. (1991). Les besoins de perfectionnement des diplômés du CPEC-PERFORMA: Comment y répondre? Sherbrooke: Université de Sherbrooke.

Everitt, B. S. (1988). Contingency Tables. J. P. Keeves (editor), Educational Research, Methodology, and Measurement: An International Handbook (first edition). Toronto: Pergamon Press (pp. 609-613).

Fédération des cégeps (1993). Cadre de référence pour une politique organisationnelle de développement des compétences. Montréal.

Fontaine, F. & Trahan, M. (1990, avril). Pour une mesure des croyances dans un contexte de formation. Mesure et évaluation en éducation, 12(4), 5-22.

Fontaine, F. (1988). Mesure critériée des croyances des futurs maîtres à l'égard de l'évaluation des apprentissages. Thèse inédite de doctorat, Faculté des sciences de l'éducation. Université de Montréal, Montréal.

Forcier, P. (1991, février). Evaluation, dites-vous? Non, "assessment". Pédagogie Collégiale, 4(3), 36-39.

Frary, R. B., Cross, L. H. & Weber, L. S. (1992). Testing and grading practices and opinions in the nineties: 1890s or 1990s? [Communication présentée dans le cadre du colloque de l'AERA à San-Francisco].

Frisbie, R. D. (1991). New Directions in evaluation: A school district perspective. Theory Into Practice, 30(1), 49.

Gadbois, L., Burelle, R., Parent, C. & Séguin, S. P. (1991). Un instrument de mesure des croyances et attitudes des enseignants à l'égard des pratiques d'évaluation formative des apprentissages de leurs élèves. Mesure et évaluation en éducation, 14(2), 5-24.

Geoffrion, P. (1992). Le groupe de discussion. B. Gauthier (éditeur), Recherche Sociale (2ième édition). Sainte-Foy: P.U.Q. 311-335.

Gong , B. & Venezky, R. (1992, janvier). Systemic change in science education: What role can classroom-based assessment play? [Communication présentée dans le cadre du colloque de l'AERA à San-Francisco].

Gravel, L. (1991). L'évaluation du rendement de l'élève dans les écoles élémentaires et secondaires de langue française de la région du Nord-Est. Rapport d'enquête. Services consultatifs de langue française, Région du Nord-Est, ministère de l'Éducation. Ontario.

Green, S. B. (1992, avril). Student and teacher judgments about lower and higher-order test items. [Communication présentée dans le cadre du colloque de L'AERA à San-Francisco].

Guba, E. G. & Lincoln, Y. S. (1991). Effective evaluation. San Francisco: Jossey-Bass pub.

Hadji, C. (1990). L'apprentissage assisté par l'évaluation: mythe ou réalité. Cahiers Pédagogiques, (no. 281), 20-23.

Howe, R. (1991). Formules pédagogiques et évaluation formative: une combinaison gagnante. Pédagogie Collégiale, 4(4), 8-13.

Howe, R. (1992). Une Typologie des champs de compétence des professeurs en évaluation des apprentissages: les catégories de Stiggins. Mesure et évaluation en éducation, 15(3), 65-76.

Howe, R. & Ménard, L. (1993, mars). Un nouveau paradigme en évaluation des apprentissages. Pédagogie Collégiale, 6(3), 36-40.

Huberman, A. M. & Miles, M. B. (1991). Analyse des données qualitatives. Recueil de nouvelles méthodes. Bruxelles: De Boeck Wesmael.

Inchauspé, P. (1992, mai). Quelle école pour demain ? Pédagogie Collégiale, 5(4), 5-10.

Jett, D. L. & Schafer, W. D. (1992, avril). Classroom teachers move to center stage in the assessment arena, ready or not. [Communication présentée dans le cadre du colloque de l'AERA à San-Francisco].

Legendre, R. (1993). Dictionnaire Actuel de l'Éducation. (2ième édition) Montréal: Guérin.

Lundeberg, M. A. & Fox, P. W. (1991). Do laboratory findings on test expectancy generalize to classroom outcomes? Review of Educational Research, 61(1), 94-106.

Marso, R. N. & Pigge, F. L. (1992). A summary of published research: Classroom teachers' knowledge and skills related to the development and use of teacher-made tests". [Communication présentée dans le cadre du colloque de L'AERA à San-Francisco].

Martineau, G. (1984). Réflexions sur l'évaluation sommative. Mesure et évaluation en éducation, 7(1), 67-70.

McMorris, R. F. & Boothroyd, R. A. (1992, avril). Tests that teachers build: an analysis of classroom tests in science and mathematics. [Communication présentée dans le cadre du colloque de L'AERA à San-Francisco].

Ménard, L. (1990). L'utilisation de l'écriture au collégial: étude descriptive. . Laval: collège Montmorency. (Rapport de recherche PAREA).

Milton, O., Pollio, H. R. & Eison, J. A. (1986). Making sense of college grades. San-Francisco: Josey-Bass.

Ministère de l'Enseignement supérieur et de la Science (1993, avril). Des collèges pour le Québec du XXIe siècle: Orientations d'avenir et mesures de renouveau. Gouvernement du Québec.

Morin, B. (1989, mai). Programme d'études et compétence des étudiants. Pédagogie Collégiale, 2(4), 32-37.

Morissette, D. (1984). La mesure et l'évaluation en enseignement. Sainte-Foy: Presses de l'Université Laval.

Morissette, D. (1993). Les examens de rendement scolaire. Sainte-Foy: Presses de l'Université Laval.

Moss, P. (1992, avril). Shifting Conceptions of validity in educational measurement: Implications for performance assessment [Communication présentée dans le cadre du colloque de L'AERA, à San-Francisco].

Nadeau, G. & Comeau, M. A. (1975). Attitudes, evaluation and exam practices of faculty in a small university. [Document présenté à la NCME, Washington].

Nava, F. & Loyd, B. H. (1992, avril). The effect of student characteristics on the grading process. [Paper presented at the annual meeting of the National Council on Measurement in Education, San-Francisco].

Noizet G. & Caverni, J. P. (1978). Psychologie de l'évaluation scolaire. Paris: PUF.

Noizet G. & Caverni J.P. (1983). Les procédures d'évaluation ont-elles leur part de responsabilité dans l'échec scolaire? Revue française de pédagogie, 62, 7-14.

Nunziati, G. (1991, mai). L'évaluation formatrice. Numéro spécial des Cahiers Pédagogiques: L'évaluation. Paris: C.R.A.P. (pp. 36-40).

Ouellet, A. (1981) Processus de recherche: une approche systémique. Québec. Presses de l'Université du Québec.

Pajares, F. M. (1992). Teachers' beliefs and educational research: Cleaning Up a Messy Construct. Review of Educational Research, 62(3), 307-332.

Paris, S. G. (1991). A developmental perspective on standardized achievement testing. Educational Researcher, 20(5), 12-20.

Popham, J. W. (1991). Appropriateness of teachers' test-preparation practices. Educational Measurement: Issues and Practice, 10(4), 12-15.

Pryor , B. W. (1992, avril). Measuring and Understanding Attitude: A more Nearly Complete Measure for Survey Research [Communication présentée dans le cadre du colloque de l'AERA à San-Francisco].

Ranjard, P. (1984). Les enseignants persécutés. Paris: Robert Jauze, éd.

Richardson, V. et. al. (1991). The relashionship between teachers' beliefs and practices in reading comprehension instruction. American educational research journal, 28(3), 559-586.

Robertson, G. J. (1991, avril). Promoting good testing practice among practitioners and students: an interdisciplinary approach. Communication présentée dans le cadre du colloque du National Council on Measurement in Education.

Robitaille, M. & Maheu, L. (1993). Les réseaux sociaux de la pratique enseignante et l'identité professionnelle: le cas du travail enseignant au collégial. Revue des Sciences de l'Education, 19(1), 87-112.

Romano, G. (1991, juin). Étudier en surface ou en profondeur. AQPC. Actes du colloque 1991 (atelier 110, p. 5).

Rose, J. S. & Medway F.J. (1981). Measurement of teachers' beliefs in their control over student outcome. Journal of Educational Research, 74(3), 185-190.

Roy, D. (1991). Étude de l'importance des connaissances de l'enseignant et de l'influence des actes professionnels d'enseignement sur l'apprentissage au collégial. Rimouski: Cégep de Rimouski. (Rapport de recherche PAREA).

Scallon, G. (1988). L'évaluation formative des apprentissages. (tome I: La réflexion); (tome II: L'instrumentation). Québec: P.U.L.

Scallon, G. (1992). L'évaluation formative: entre la docimologie et la didactique. D. Laveault (éditeur), Les pratiques d'évaluation en éducation, (pp. 23-34). Montréal: Les éditions de l'ADMEE.

Schafer, W. D. & Lissitz, R. W. (1987). Measurement training for school personnel: Recommandations and reality. Journal of Teacher Education, 38(3), 57-63.

Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. Educational Measurement: Issues and Practice, 10(1), 3-6.

Shavelson, R. J. & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. Review of Educational Research, 51(4), 455-498.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. Journal of Educational Measurement, 22(4), 271-286.

Stiggins, R. J., Frisbie, D. A. & Griswold, P. A. (1989). Inside High-School grading practices: Building a research agenda. Educational Measurement: Issues and Practice, 8(2), 5-14.

Stiggins, R. J. (1991, mars). Relevant classroom assessment training for teachers. Educational Measurement: Issues and Practice, 10(1), 7-12.

Stiggins R.J., Conklin, N., & et al. (1992). In teachers' hands: Investigating the practices of classroom assessment. Albany, NY.: SUNY Press.

Tabachnick, B. G. & Fidell, L. S. (1989). Using multivariate statistics. New-York: Harper-Collins.

Terwilliger, J. S. (1977). Assigning grades: Philosophical issues and practical recommendations. Journal of research and development in education, 10(3), 21-39.

Terwilliger, J. R. (1989). Classroom Standard Setting and Grading Practices. Educational Measurement: issues and practice, 8(2), 15-19.

Tomlinson, T., & Cross, C. (1991, septembre). Student Effort: The Key to Higher Standards. Educational Leadership, 49(1), 69-73.

Tousignant, R. & Morissette, D. (1990). Les principes de la mesure et de l'évaluation des apprentissages. Boucherville: Gaetan Morin, éd.

Trudel, R. & Antonius, R. (1991). Méthodes quantitatives appliquées aux sciences humaines. Montréal: Centre éducatif et culturel.

Tyler, R. W. (1986). Changing concepts of educational evaluation. International Journal of Educational Research, 10, 11-35.

Van der Maren, J. (1990, mai). Méthodes de Recherche en Education. Exposé critique à l'intention des utilisateurs des résultats de la recherche et des chercheurs en éducation. Montréal: La librairie de l'Université de Montréal.

Wilk, R. E. & Carey, L. M. (1983). Grading student progress: teachers' opinions vs experts' advice. [Document présenté à l'AERA à Montréal] (15 p.).

Wolf, R. M. (1988). Questionnaires. J. P. Keeves (editor), Educational Research, Methodology, and Measurement: An International Handbook (first edition), (pp. 478-482). Toronto: Pergamon.

Wood, P. H. (1990). Grading and evaluation pratices and policies of school teachers. [Document présenté au colloque du National Council on Measurement in Education, Boston] (58 p.).

ANNEXE 1
LE QUESTIONNAIRE

Croyances et pratiques en évaluation des apprentissages

Laval, le 31 mars 1992

Madame,
Monsieur,

Certains chercheurs affirment que l'évaluation des apprentissages représente une partie importante de la tâche des enseignants et des enseignantes. Pourtant, nous savons bien peu de choses sur la façon dont ceux-ci évaluent et notent leurs étudiants ainsi que sur les difficultés qu'ils rencontrent.

Nous sommes, respectivement, conseiller pédagogique et enseignante au Collège Montmorency. Nous réalisons une recherche subventionnée par le programme PAREA de la DGEC. Cette recherche porte sur les opinions et les pratiques des enseignants et des enseignantes à l'égard de l'évaluation des apprentissages des étudiants. Elle vise à circonscrire la situation actuelle de l'évaluation afin de mieux connaître les problèmes qu'elle pose aux enseignants et de mettre en lumière des activités de perfectionnement adaptées à leurs besoins.

Nous sollicitons, ici, votre participation à cette recherche, car elle est indispensable à une meilleure compréhension du problème de l'évaluation au collégial.

Nous tenons à vous assurer que toutes vos réponses demeureront confidentielles et qu'elles ne seront utilisées qu'à des fins statistiques.

Nous vous prions donc de compléter ce questionnaire et de le glisser dans l'enveloppe ci-jointe. **Veuillez nous retourner le tout, par courrier, d'ici le 21 avril 1992.** Nous vous remercions à l'avance de votre collaboration.

Robert Howe et Louise Ménard
Collège Montmorency
475 Boul. de l'Avenir
Laval, H7N 5H9

Comment remplir ce questionnaire

Si vous enseignez plus d'un cours, il vous sera plus facile de répondre au questionnaire en choisissant arbitrairement l'un de ces groupes-cours et en vous référant toujours à ce même groupe-cours pour toutes les questions.

Si vous enseignez principalement à l'extérieur des murs de votre Collège (stage, industrie, clinique, etc.), nous vous demandons de vous reporter au plus récent cours que vous avez donné dans les locaux de votre institution. D'ailleurs, c'est l'ensemble des activités d'évaluation ayant lieu ***sur le campus de votre Collège*** auxquelles nous faisons référence lorsque nous utilisons dans ce questionnaire l'expression "salle de classe".

Nous vous recommandons de répondre en une seule séance. Indiquez vos choix de réponse en **cochant le chiffre** correspondant, sur les échelles de jugement ou sur les choix proposés. Il n'y a pas de "bonne réponse". C'est votre opinion seule qui compte.

Dans ce questionnaire, nous convenons des distinctions suivantes :

- l'évaluation se réfère à toute activité qui permet à l'enseignant de porter un jugement sur les apprentissages des étudiants.
- l'évaluation formative s'intéresse à la progression de l'étudiant, pendant la formation, afin d'ajuster l'enseignement ou de corriger les difficultés d'apprentissage.
- l'évaluation sommative mesure la somme des apprentissages en vue d'une décision de promotion.

1) Cette session, l'ensemble de mes instruments d'évaluation mesureront (cochez une réponse) :

- ❑ 1- tous les éléments du contenu du cours
- ❑ 2- un échantillon du contenu du cours
- ❑ 3- seulement les éléments vraiment essentiels du contenu

2) Mes instruments d'évaluation visent la mesure de :	Toujours	Souvent	Parfois	Rarement	Jamais
a- la connaissance d'un contenu disciplinaire	①	②	③	④	⑤
b- la compréhension de la matière	①	②	③	④	⑤
c- la mise en application des procédures explicitement enseignées (savoir-faire)	①	②	③	④	⑤
d- l'habileté à résoudre des problèmes complexes	①	②	③	④	⑤
e- la capacité d'analyse et de synthèse	①	②	③	④	⑤

	Toujours	Souvent	Parfois	Rarement	Jamais
f- la pensée critique	①	②	③	④	⑤
g- les attitudes (savoir-être)	①	②	③	④	⑤

3) Dans ce cours, en plus d'attribuer des points pour les objectifs liés à la matière enseignée, j'en accorde pour :

a- la motivation et l'effort dans l'apprentissage	❑ 1. oui	❑ 2. non
b- la présence assidue aux cours	❑ 1. oui	❑ 2. non
c- l'exécution de devoirs ou d'exercises à la maison ou en classe	❑ 1. oui	❑ 2. non
d- les progrès réalisés au cours de la session (évolution)	❑ 1. oui	❑ 2. non
e- la participation active de l'étudiant en classe	❑ 1. oui	❑ 2. non
f- la créativité, l'originalité	❑ 1. oui	❑ 2. non
g- le respect de mes exigences de présentation des travaux	❑ 1. oui	❑ 2. non

	Toujours	Souvent	Parfois	Rarement	Jamais
4) Les travaux et les examens que je soumets à mes étudiants contribuent au calcul de la note inscrite au bulletin.	①	②	③	④	⑤

5) Combien d'heures, sur un format de cours de ***45 heures***, vos étudiants sont-ils occupés à des ***activités d'évaluation*** de toute nature (exclure les travaux faits à l'extérieur du collège, tels les travaux réalisés à la maison) ? (cochez une réponse)

- ❑ 1- moins de 3 heures
- ❑ 2- de 3 à 5 heures
- ❑ 3- de 6 à 8 heures
- ❑ 4- de 9 à 12 heures
- ❑ 5- de 13 à 18 heures
- ❑ 6- plus de 18 heures

6) Voici une série d'affirmations. Indiquez si vous êtes personnellement...

1. Très en accord
2. En accord
3. En désaccord
4. Très en désaccord
5. Ne sais pas

Cocher votre choix de réponse	Très en accord	En accord	En désaccord	Très en désaccord	Ne sais pas
a- La meilleure façon de faire travailler les étudiants est de donner des points pour leur travail.	①	②	③	④	⑤
b- Le fait d'être évalués encourage les étudiants à fournir plus d'effort à l'étude et au travail scolaire.	①	②	③	④	⑤
c- L'évaluation sert à identifier les forces et les faiblesses de l'étudiant en fonction des apprentissages à faire.	①	②	③	④	⑤
d- Suite aux résultats d'une évaluation des apprentissages, les enseignants devraient être disposés à réajuster leur contenu de cours.	①	②	③	④	⑤
e- L'évaluation sert à vérifier ce que les étudiants ont appris.	①	②	③	④	⑤
f- Si je le pouvais, je n'évaluerais pas les étudiants.	①	②	③	④	⑤
g- Les étudiants travaillent de façon plus régulière lorsque l'évaluation est fréquente.	①	②	③	④	⑤
h- L'évaluation peut aider les étudiants à apprendre.	①	②	③	④	⑤
i- L'évaluation des apprentissages des étudiants sert principalement à satisfaire aux exigences administratives du Collège.	①	②	③	④	⑤
j- Au collégial, il n'est pas nécessaire de faire de l'évaluation formative.	①	②	③	④	⑤
k- L'évaluation formative, ce sont de petites évaluations qui servent à préparer à l'évaluation sommative.	①	②	③	④	⑤
l- Les examens devraient porter sur tout ce qui est dit en classe.	①	②	③	④	⑤
m- Il **n'est pas** correct de poser des questions d'examen au-delà de la matière enseignée.	①	②	③	④	⑤
n- Il **n'est pas** utile que l'évaluation porte sur toute la matière enseignée.	①	②	③	④	⑤
o- Il est impossible, au collégial, d'évaluer autre chose que la connaissance de la matière.	①	②	③	④	⑤
p- Il est pratiquement impossible d'évaluer les attitudes (le savoir-être).	①	②	③	④	⑤
q- Il est plus important d'évaluer la compréhension de la matière que la connaissance des faits.	①	②	③	④	⑤
r- Au collégial, on devrait surtout évaluer les habiletés intellectuelles supérieures (analyse, synthèse, résolution de problèmes complexes).	①	②	③	④	⑤

7) Lorsque je corrige un examen ou un travail, je me base généralement sur des critères d'évaluation déterminés : (cochez une réponse)

- ❑ 1- à l'avance.
- ❑ 2- juste avant le début de la correction
- ❑ 3- au cours de la correction
- ❑ 4- après une première lecture des copies des étudiants

	Toujours	Souvent	Parfois	Rarement	Jamais
8) Lorsque je rédige un instrument de mesure, je commence par me faire un plan détaillé de ce que je veux évaluer.	①	②	③	④	⑤
9) Lorsque j'évalue les apprentissages, je donne les critères de correction aux étudiants à l'avance.	①	②	③	④	⑤

10) **Dans les deux dernières semaines**, pour ce groupe-cours, combien de temps avez-vous pris pour la **préparation** (planification, rédaction) des moyens d'évaluation contribuant au calcul de la note finale ? (cochez une réponse)

- ❑ 1- moins de 3 heures
- ❑ 2- de 3 à 6 heures
- ❑ 3- de 7 à 9 heures
- ❑ 4- de 10 à 12 heures
- ❑ 5- plus de 12 heures

11) Dans ce cours, j'évalue les apprentissages des étudiants pour des fins de notation au bulletin : (cochez une réponse)

- ❑ 1- à toutes les semaines
- ❑ 2- aux 2 semaines
- ❑ 3- aux 3 semaines
- ❑ 4- aux 5 semaines ou plus rarement

12) Utilisez-vous l'examen pour effectuer l'évaluation de vos étudiants ?

- ❑ 1. oui
- ❑ 2. non

Si vous utilisez l'examen, répondez aux questions 13 à 29
Si vous n'utilisez *pas* l'examen, passez à la question 30

	Toujours	Souvent	Parfois	Rarement	Jamais
13) Je fais examiner mes questions par un ou des collègues avant d'administrer un examen.	①	②	③	④	⑤
14) Après la correction d'un examen, je fais l'analyse de mes questions d'examen (clarté, compréhension, difficulté, plausibilité des choix de réponse s'il y a lieu, etc.).	①	②	③	④	⑤
15) Je prépare la plupart de mes questions d'examen à partir de questions déjà existantes.	①	②	③	④	⑤
16) Je prépare la plupart de mes questions d'examen en relisant le contenu de mon cours.	①	②	③	④	⑤
17) Je prépare la plupart de mes questions d'examen à partir des objectifs de mon cours.	①	②	③	④	⑤
18) Lorsque je donne un examen, je préviens les étudiants, à l'avance, de ce sur quoi portera l'examen.	①	②	③	④	⑤
19) Pour aider les étudiants à se préparer à répondre aux questions d'examen, je...					
a- fais faire ou propose des exercices avec des questions équivalentes.	①	②	③	④	⑤
b- fais rédiger une partie des questions d'examen par les étudiants eux-mêmes.	①	②	③	④	⑤
c- fais des démonstrations et/ou des revisions de la matière.	①	②	③	④	⑤
20) Pendant un examen, je présente mes questions d'examen de la façon suivante :					
a- j'écris les questions au tableau ou je les projette sur acétate.	①	②	③	④	⑤
b- je les imprime et en distribue un exemplaire à chaque étudiant.	①	②	③	④	⑤
c- j'en fais la lecture à haute voix et les étudiants les prennent en note.	①	②	③	④	⑤
21) Dans mes examens, j'utilise les genres d'items suivants :					
a- questions à développement court	①	②	③	④	⑤
b- questions à développement long	①	②	③	④	⑤
c- questions objectives, à choix multiples	①	②	③	④	⑤
22) Lorsque je remets les résultats d'un examen à mes étudiants, je leur demande leur opinion sur la formulation des items du questionnaire.	①	②	③	④	⑤
23) Il m'arrive de donner un examen plutôt facile pour compenser un examen antérieur difficile.	①	②	③	④	⑤

	Toujours	Souvent	Parfois	Rarement	Jamais	Ne s'applique pas
24) Dans mes examens à développement, je propose des questions «au choix» (ex : «répondre à 3 questions parmi un choix de 8»).	①	②	③	④	⑤	⑥
25) Dans mes examens à développement, je permets aux étudiants de répondre «à la maison» et de remettre leurs réponses lors de la prochaine rencontre.	①	②	③	④	⑤	⑥
26) Dans mes examens à développement, je permets aux étudiants d'avoir en main toute documentation utile.	①	②	③	④	⑤	⑥
27) Pour compenser la pratique du choix au hasard (guessing), lors des examens objectifs (à choix multiples), j'utilise un système de pénalité en enlevant des points dans certaines conditions.	①	②	③	④	⑤	⑥
28) Lorsque j'évalue le travail fait par une équipe d'étudiants, j'accorde la même note à tous les membres de l'équipe.	①	②	③	④	⑤	⑥

	Toujours	Souvent	Parfois	Rarement	Jamais
29) Je conserve mes questions d'examen en vue de les réutiliser.	①	②	③	④	⑤

30) Utilisez-vous le travail écrit pour effectuer l'évaluation de vos étudiants ?

- ❑ 1. oui
- ❑ 2. non

Si vous utilisez le travail écrit, répondez aux questions 31 et 32
Si vous n'utilisez *pas* le travail écrit, passez à la question 33

	Toujours	Souvent	Parfois	Rarement	Jamais
31) Je fais examiner mes consignes de travaux écrits par un ou des collègues avant de les présenter aux étudiants.	①	②	③	④	⑤
32) Pour aider les étudiants à exécuter les travaux écrits qui contribuent au calcul de la note finale,					
a- j'explique la tâche oralement	①	②	③	④	⑤
b- je remets une feuille explicative	①	②	③	④	⑤
c- je prévois des étapes de rétroaction corrective avant la date de remise du travail	①	②	③	④	⑤

33) Voici une série d'affirmations. Indiquez si vous êtes personnellement...

1. Très en accord
2. En accord
3. En désaccord
4. Très en désaccord
5. Ne sais pas

Cocher votre choix de réponse	Très en accord	En accord	En désaccord	Très en désaccord	Ne sais pas
a- Après la passation d'un examen, il est utile de faire l'analyse des questions pour en évaluer la qualité.	①	②	③	④	⑤
b- Les questions d'examen ont davantage de valeur lorsqu'elles sont lues par un second spécialiste de la matière.	①	②	③	④	⑤
c- On s'assure davantage de la clarté des consignes des travaux écrits en les faisant examiner par un second spécialiste de matière.	①	②	③	④	⑤
d- Tous les étudiants d'un même cours devraient être évalués à partir des mêmes critères.	①	②	③	④	⑤
e- Il est pratiquement impossible de faire en sorte que les résultats des évaluations représentent bien les apprentissages des étudiants.	①	②	③	④	⑤
f- Il faut faire plusieurs évaluations pour pouvoir se fier davantage à l'ensemble des résultats obtenus.	①	②	③	④	⑤
g- Mes moyens d'évaluation me disent ce que je veux savoir.	①	②	③	④	⑤
h- Beaucoup d'enseignants manquent de rigueur à l'égard de l'évaluation.	①	②	③	④	⑤
i- Il est nécessaire d'établir ses critères d'évaluation avant la correction.	①	②	③	④	⑤
j- Les notes devraient refléter le nombre ou le pourcentage des objectifs maîtrisés par les étudiants.	①	②	③	④	⑤
k- Seules les questions à développement peuvent mesurer les apprentissages de haut niveau cognitif (application, analyse, synthèse).	①	②	③	④	⑤
l- Les examens devraient être, de préférence, à livre ouvert.	①	②	③	④	⑤
m- Les évaluations formatives devraient surtout être des auto-évaluations.	①	②	③	④	⑤
n- Les évaluations sommatives devraient surtout être des auto-évaluations.	①	②	③	④	⑤
o- Les questions à choix multiples mesurent principalement les connaissances mémorisées.	①	②	③	④	⑤
p- Des examens-synthèse de fin de session devraient être obligatoires dans presque toutes les disciplines.	①	②	③	④	⑤
q- La rédaction de travaux écrits ou la réalisation de projets sont les meilleurs moyens d'évaluation.	①	②	③	④	⑤
r- Les examens avec questions à choix multiples peuvent mesurer les apprentissages de haut niveau cognitif (application, analyse, synthèse).	①	②	③	④	⑤
s- Les étudiants répondent trop souvent au hasard à des questions à choix multiples.	①	②	③	④	⑤

34) Utilisez-vous les approches suivantes pour effectuer l'évaluation de l'apprentissage de vos étudiants ?

	Toujours	Souvent	Parfois	Rarement	Jamais
a- devoirs ou exercices à la maison	①	②	③	④	⑤
b- dossier anecdotique, journal de bord	①	②	③	④	⑤
c- étude de cas	①	②	③	④	⑤
d- observation directe avec grille d'appréciation et/ou liste de vérification (check-list)	①	②	③	④	⑤
e- mini-test ou jeux questionnaire	①	②	③	④	⑤
f- rapport ou compte-rendu d'activités ou de laboratoire	①	②	③	④	⑤
g travail écrit, recherche	①	②	③	④	⑤
h- questions orales	①	②	③	④	⑤
i- projet thématique	①	②	③	④	⑤
j- résolution de problème	①	②	③	④	⑤
k- démonstration de savoirs-faire	①	②	③	④	⑤
l- exercices en classe	①	②	③	④	⑤
m- examen surprise	①	②	③	④	⑤
n- examen écrit	①	②	③	④	⑤
o- l'auto-évaluation	①	②	③	④	⑤
p- l'évaluation par les pairs	①	②	③	④	⑤
q- schéma ou réseau de concepts	①	②	③	④	⑤

☞ ***Pour les questions 35 à 41, répondez en utilisant la lettre (a à q) correspondant à l'approche choisie.***

35) Laquelle de ces approches vous permet ***le mieux*** de savoir si vos étudiants, à la fin de la session, ont atteint les objectifs visés ? (un seul choix) ☐

36) Laquelle vous informe ***le mieux*** des progrès de l'étudiant ? (un seul choix) ☐

37) Laquelle de ces approches utiliseriez-vous ***davantage*** si vous la connaissiez mieux ? (un seul choix) ☐

38) Laquelle de ces approches utiliseriez-vous ***davantage*** si vous aviez plus de temps ? (un seul choix) ☐

39) Laquelle de ces approches vous semble ***la plus utile*** pour aider vos étudiants à apprendre ? (un seul choix) ☐

40) Laquelle de ces approches semble ***la plus utile*** pour accumuler des notes en vue du bulletin de fin de session ? (un seul choix) ☐

41) Lesquelles de ces approches utilisez-vous pour évaluer vos étudiants pendant leur apprentissage, dans un but formatif, sans donner de note allant au bulletin ? ☐

42) Dans mon département, le choix des types d'instruments d'évaluation est fait par : (cochez une réponse)

- ❑ 1- moi seul.
- ❑ 2- moi et mes collègues donnant le même cours.
- ❑ 3- moi et mon département.
- ❑ 4- moi et mes étudiants.

43) Le choix de ce sur quoi doit porter mon évaluation est généralement fait par : (cochez une réponse)

- ❑ 1- moi seul
- ❑ 2- moi et mes collègues donnant le même cours
- ❑ 3- moi et mon département
- ❑ 4- moi et mes étudiants

44) Voici une série d'affirmations. Indiquez si vous êtes personnellement...

1. Très en accord
2. En accord
3. En désaccord
4. Très en désaccord
5. Ne sais pas

Cocher votre choix de réponse	*Très en accord*	*En accord*	*En désaccord*	*Très en désaccord*	*Ne sais pas*
a- Les étudiants qui travaillent fort méritent d'avoir des notes plus élevées que ceux qui ne le font pas.	①	②	③	④	⑤
b- De mauvais résultats d'évaluation ont un effet démotivant sur les étudiants.	①	②	③	④	⑤
c- Il faut s'abstenir d'augmenter la note d'un étudiant pour l'encourager ou le pousser à travailler davantage.	①	②	③	④	⑤
d- Lors de la correction des questions à développement ou des travaux écrits, les enseignants sont influencés par la connaissance de l'identité du répondant.	①	②	③	④	⑤
e- C'est la conception qu'on se fait de l'enseignement qui influence le plus nos choix de stratégies d'évaluation.	①	②	③	④	⑤

45) Il m'arrive de **diminuer** une note finale transmise au bulletin lorsque je veux pénaliser :

a- des absences non motivées ❑ 1. oui ❑ 2. non

b- le peu d'effort fourni par l'étudiant ❑ 1. oui ❑ 2. non

46) Il m'arrive d'**augmenter** une note finale transmise au bulletin lorsque je veux récompenser :

a- la motivation et l'effort dans l'apprentissage ❑ 1. oui ❑ 2. non

b- la présence assidue aux cours ❑ 1. oui ❑ 2. non

c- l'exécution des devoirs ou des travaux à la maison ou en classe ❑ 1. oui ❑ 2. non

d- la réalisation de travaux supplémentaires ❑ 1. oui ❑ 2. non

e- les progrès réalisés au cours de la session (évolution) ❑ 1. oui ❑ 2. non

f- la participation active de l'étudiant en classe ❑ 1. oui ❑ 2. non

g- la créativité, l'originalité ❑ 1. oui ❑ 2. non

	Toujours	Souvent	Parfois	Rarement	Jamais
47) J'ai tendance à encourager un étudiant faible en accordant une note de passage lorsque ses efforts me semblent appréciables.	①	②	③	④	⑤
48) Il m'arrive de modifier les notes de tous mes élèves pour ajuster la moyenne de classe.	①	②	③	④	⑤

49) Voici une série d'affirmations. Indiquez si vous êtes personnellement...

1. Très en accord
2. En accord
3. En désaccord
4. Très en désaccord
5. Ne sais pas

Cocher votre choix de réponse	*Très en accord*	*En accord*	*En désaccord*	*Très en désaccord*	*Ne sais pas*
a- La note moyenne de la classe est un reflet direct de la qualité de l'enseignement reçu.	①	②	③	④	⑤
b- Dans un groupe, la distribution des notes devrait suivre la courbe normale, en forme de cloche : peu de faibles, beaucoup de moyens et peu de forts.	①	②	③	④	⑤
c- La notation nuit à l'enseignement.	①	②	③	④	⑤
d- Plusieurs enseignants accordent une note de passage à des étudiants qui ne la méritent pas.	①	②	③	④	⑤
e- Les enseignants devraient généralement commenter par écrit les productions écrites de leurs étudiants.	①	②	③	④	⑤
f- Il faut éviter de faire des évaluations qui impliquent le jugement personnel, la subjectivité de l'enseignant.	①	②	③	④	⑤
g- Les résultats d'une évaluation formative ne doivent pas servir au calcul de la note inscrite au bulletin.	①	②	③	④	⑤
h- La note, ça représente la paye de l'étudiant.	①	②	③	④	⑤
i- La plupart des étudiants lisent les commentaires que leurs professeurs écrivent sur leurs feuilles-réponses ou sur leurs travaux écrits.	①	②	③	④	⑤

	Toujours	Souvent	Parfois	Rarement	Jamais
50) La note finale que j'inscris au bulletin de l'étudiant contient, entre autres, des notes obtenues dans des évaluations formatives.	①	②	③	④	⑤
51) Dans la détermination des notes finales, je tiens compte de la politique d'évaluation des apprentissages adoptée par mon département ou par mon collège.	①	②	③	④	⑤
52) Lorsque je retourne un examen ou un travail aux étudiants, je donne à chacun une rétroaction détaillée sur les points à améliorer.	①	②	③	④	⑤
53) Après un examen ou un travail, je prends les moyens nécessaires pour que les étudiants puissent améliorer leurs points faibles.	①	②	③	④	⑤

54) Pour ce cours, je retourne généralement les examens et travaux corrigés aux étudiants après : (cochez une réponse)

- ❑ 1- une semaine
- ❑ 2- deux semaines
- ❑ 3- trois semaines
- ❑ 4- quatre semaines ou plus

Formation et perfectionnement

55) Quelle préparation avez-vous reçue en matière d'**évaluation des apprentissages** ? (vous pouvez cocher plus d'une réponse)

- ❑ 1- Aucune
- ❑ 2- Au moins un cours au baccalauréat
- ❑ 3- Au moins un cours au CPEC-Performa
- ❑ 4- Au moins un cours à la maîtrise
- ❑ 5- Au moins une journée pédagogique
- ❑ 6- Au moins une conférence ou un atelier
- ❑ 7- Autodidaxie

56) Comment jugez-vous vos compétences dans les domaines suivants ?	Très bonne	Bonne	Passable	Faible	Nulle
a- Évaluation formative	①	②	③	④	⑤
b- Évaluation sommative	①	②	③	④	⑤

57) Si on organisait des activités de perfectionnement en mesure et évaluation, quels sujets, parmi les suivants, vous intéresseraient davantage ? (vous pouvez cocher plusieurs sujets)

- ❑ 1- Rédaction d'objectifs
- ❑ 2- Construction d'examens
- ❑ 3- Techniques d'observation
- ❑ 4- Évaluation formative
- ❑ 5- Évaluation fondée sur les compétences à acquérir
- ❑ 6- Évaluation des attitudes
- ❑ 7- Planification de l'évaluation
- ❑ 8- L'attribution des notes
- ❑ 9- La correction (annotation)
- ❑ 10- aucune
- ❑ 11- autre (précisez) _____

58) Quelles ressources, parmi les suivantes, vous seraient utiles ?

- ❑ 1- Des modèles de consignes pour travaux écrits
- ❑ 2- Une banque de questions d'examens
- ❑ 3- Des examens de fin de session déjà construits
- ❑ 4- Des examens modèles
- ❑ 5- Un système informatisé de correction d'examens à choix multiples
- ❑ 6- Un système informatisé d'analyse d'examens
- ❑ 7- aucune
- ❑ 8- Autre (précisez) _____

Informations générales

59) Nom du collège : _____

60) Quels sont les trois **premiers chiffres des cours** que vous enseignez habituellement : _____

61) Sexe : ❑ 1. féminin ❑ 2. masculin

62) Age : ☐

63) Années d'expérience reconnues dans l'enseignement : ☐

64) Années de scolarité reconnues pour fins de rémunération : ☐

65) Niveau de scolarité complété :

- ❑ 1- DEC ou l'équivalent
- ❑ 2- Certificat de premier cycle
- ❑ 3- Baccalauréat
- ❑ 4- Maîtrise
- ❑ 5- Doctorat
- ❑ 6- Autre, précisez _____

66) Quelle(s) **formule(s) pédagogique(s)** utilisez-vous pour ce cours ? (vous pouvez cocher plusieurs items)

- ❑ 1- L'exposé informel en classe
- ❑ 2- L'exposé magistral formel
- ❑ 3- Le laboratoire
- ❑ 4- Le travail en équipe
- ❑ 5- Le travail individuel (exercices individuels)
- ❑ 6- Le séminaire ou groupe de discussion
- ❑ 7- Le stage

67) Commentaires _____

_____

_____

_____

_____

ANNEXE 2

VALIDATION DU QUESTIONNAIRE

Sources documentaires pour les questions du questionnaire

Les sources primaires des énoncés de croyances ou de pratiques

Toutes les questions présentes dans le questionnaire ont été extraites ou inspirées de sources de différentes natures. La grande majorité proviennent de notre revue de la littérature et ont été inspirées des textes ou des questionnaires utilisés dans des recherches semblables. Certains énoncés proviennent de notre expérience professionnelle dans un cégep: observation, anecdotes, lecture de plans de cours.

Sources documentaires

La plupart de nos énoncés de questions ont été inspirés des sources suivantes:

Agnew (1985)	Milton(1986)
Anderson (1989)	Nadeau et Comeau (1975)
Bloom (1979)	Noizet et Caverni (1983)
Chase (1984)	Stiggins (1989)
Fontaine (1988)	Terwilliger (1989)
Gravel (1991)	Tomlinson (1991)
Conseil des collèges (1985)	Wilk (1983)
Martineau (1984)	Wood (1990)

Tous les énoncés ont été transcrits individuellement et ont d'abord été classés dans l'un ou l'autre groupe: énoncés de croyances ou énoncés de pratiques. Puis, les énoncés ont été identifiés à l'une ou l'autre des six catégories de Stiggins. Avant l'élagage de tous ces énoncés, il a été observé que certaines catégories d'énoncés étaient nettement plus alimentées que d'autres. On a noté une abondance d'énoncés appartenant aux catégories:

C6 Croyances relatives à la notation;

P4 Pratiques relatives à l'instrumentation en évaluation.

Par contre, il y avait carence d'énoncés dans les catégories suivantes:

C3 Croyances relatives aux qualités de l'évaluation;

P1 Pratiques relatives à l'utilisation de l'évaluation en salle de classe;

P2 Pratiques relatives aux cibles de l'évaluation.

Ces déséquilibres peuvent, croyons-nous, illustrer les zones qui ont été privilégiées et celles qui ont été négligées dans les discours et recherches en évaluation des apprentissages.

Le jury de validation du questionnaire

Liste des consultants-experts	Fonction
Madame Cécile d'Amour	Groupe de Recherche-Action, programme Performa, université de Sherbrooke.
Madame France Fontaine	Spécialiste en mesure et évaluation et adjointe à la Doyenne des études de premier cycle à l'U.Q.A.M.
Madame Louise Landry	Professeure au cégep de Rosemont et membre de la Commission d'évaluation du Conseil des collèges.
Madame Hélène Lavoie	Professionnelle responsable du programme d'aide à la recherche sur l'enseignement et l'apprentissage à la DGEC.
Monsieur Jean-Paul Michaud	Conseiller pédagogique au cégep de Rivière-du-Loup.
Monsieur Dominique Morissette	Professeur en mesure et évaluation à l'Université Laval et auteur de plusieurs livres en mesure et évaluation.
Monsieur Philippe Ricard	Expert-conseil en enquête et sondage et directeur général du Bureau d'études sociographiques.

DOCUMENT 1

ETAPE DE PRÉ-VALIDATION PAR JURY D'EXPERTS

DESCRIPTION DU CONTEXTE

Le questionnaire que nous préparons est destiné à un échantillon de professeurs en exercice, dans les cégeps du Québec, pendant la session d'hiver 1992. L'échantillon sera composé de professeurs représentant l'ensemble des disciplines enseignées, tant au secteur général qu'au secteur professionnel.

Par ce questionnaire, nous cherchons à connaître:
-les pratiques de ces professeurs en matière d'évaluation des apprentissages en salle de classe;
-les croyances sur lesquelles reposeraient ces pratiques.

Pour les questions portant sur les croyances, les professeurs seront invités à indiquer leur degré d'accord ou de désaccord à chaque énoncé en utilisant l'échelle suivante:

0-NE SAIS PAS
1-TOUT À FAIT EN DÉSACCORD
2-PLUTÔT EN DÉSACCORD
3-PLUTÔT EN ACCORD
4-TOUT À FAIT EN ACCORD

Quant aux questions sur les pratiques, plusieurs proposeront des CHOIX DE RÉPONSE (cela sera évident pour vous dans les énoncés) alors que les autres permettront de répondre aux choix suivants: TOUJOURS-SOUVENT-RAREMENT-JAMAIS.

Les réponses à ce questionnaire, une fois traitées et analysées, nous permettront d'explorer la possibilité de liens entre les croyances et les pratiques. Pour ce faire, les résultats du questionnaire seront proposés, à l'automne 1992, à un autre échantillon plus petit de professeurs. Ces personnes seront interviewées individuellement et seront invitées à évaluer les données et à proposer leur interprétation personnelle sur les croyances et pratiques observées, sur les liens qu'ils perçoivent entre les croyances et les pratiques ainsi que sur les besoins de perfectionnement des professeurs en évaluation des apprentissages.

CONSIGNES AU JURY

Tous les énoncés qui suivent vous sont présentés sans les échelles de jugement ou de choix de réponse qui seront bien sûr fournis dans la version finale du questionnaire. Ces énoncés, pour faciliter votre tâche, sont regroupés et présentés tels qu'ils ont été catalogués et identifiés depuis le début: par catégorie et par type (croyances ou pratiques).

Ce qui suit a déjà fait l'objet d'un élagage sévère car nous avions recueilli plus de deux cents énoncés de questions avant Noël. C'est dans la documentation spécialisée que nous avons extrait la majorité des énoncés. Les autres nous viennent surtout des entrevues de type semi-dirigé que nous avons tenues auprès de 12 professeurs du Collège Montmorency à la fin de l'automne. Tel que prévu dans notre mémoire de présentation de ce projet de recherche, nous avons donné priorité (mais non exclusivité) aux énoncés portant sur le processus plutôt que sur le produit de l'évaluation.

Pour regrouper nos énoncés, nous avons choisi un système de catégories basé sur un texte récent de R. Stiggins (1991). Cet auteur décrit, en six dimensions, les compétences nécessaires aux professeurs en matière de mesure et évaluation.. Les thèmes de ces dimensions constituent nos titres de catégories et les descriptions que nous fournissons ci-après pour chaque catégorie sont, en fait, des résumés d'une traduction libre que nous avons réalisée du texte de Stiggins. Afin de vous faciliter une vue d'ensemble de nos catégories, nous vous fournissons cette adaptation que nous avons faite de la partie de l'article de Stiggins traitant des compétences en évaluation. C'est le DOCUMENT 3.

Les jugements et opinions que nous vous demandons sont les suivants:

-clarté et pertinence de chacun des énoncés de questions;
-pertinence des catégories;
-commentaires et suggestions.

Dans le document qui suit, appelé DOCUMENT 2, les énoncés sont accompagnés de deux colonnes, intitulées "clarté" et "pertinence". Dans la première, vous indiquez votre jugement sur le degré de clarté de chaque énoncé et, dans la seconde, votre jugement sur le degré de pertinence de l'énoncé. Utilisez les chiffres de l'échelle suivante:

4- TOUT À FAIT CLAIR (OU PERTINENT)
3- PLUTÔT CLAIR (OU PERTINENT)
2- MANQUE PLUTÔT DE CLARTÉ (OU DE PERTINENCE)
1- MANQUE TOUT À FAIT DE CLARTÉ (OU DE PERTINENCE)

Voici un exemple:

	clarté	pertinence
"L'évaluation, c'est seulement une affaire de notes"	2	3

La cote 2 indique ici que l'énoncé serait considéré par le répondant comme manquant plutôt de clarté alors que la cote 3 révèle que l'énoncé est estimé comme plutôt pertinent.

Après chaque série, un espace de réponse libre vous est proposé pour que vous puissiez y indiquer vos commentaires sur la pertinence de chaque catégorie ou vos commentaires ou suggestions de rédaction / modification au sujet des énoncés. Nous vous prions de vous référer au numéro d'énoncé, s'il y a lieu. Si l'espace manquait, n'hésitez pas à utiliser des feuilles supplémentaires.

Pour toute autre explication, nous vous invitons à nous téléphoner, à frais virés s'il y a lieu, à la maison (514-664-1833) ou au bureau (514-667-5100, poste 189).

S'il vous plaît, nous apprécierions que vous nous retourniez vos jugements, sur le DOCUMENT 2, avant le 3 février.

HOWE / MÉNARD, JANVIER 1992

DOCUMENT 2

IDENTIFICATION DU JUGE:_____

CROYANCES ET PRATIQUES EN ÉVALUATION

Robert Howe et Louise Ménard Janvier 1992

CATÉGORIE 1: L'utilisation de l'évaluation en classe

Cette catégorie se réfère aux buts pour lesquels l'enseignant évalue. L'enseignant se sert de l'évaluation, en général, pour appuyer ses décisions, pour guider son enseignement, pour gérer la classe, pour contrôler les comportements des étudiants et les motiver à travailler.

ÉNONCÉS DE CROYANCES

		CLARTÉ	PERTINENCE
1 -	La meilleure façon de faire travailler mes étudiants est de noter leur travail (donner des points).		
2 -	Le fait d'être évalués encourage les étudiants à fournir plus d'effort à l'étude.		
3 -	Il faut évaluer souvent pour aider les étudiants à identifier rapidement leurs faiblesses.		
4 -	L'évaluation sert à identifier les forces et les faiblesses de l'étudiant en fonction des apprentissages à faire.		
5 -	L'évaluation doit servir à classer les étudiants les uns par rapport aux autres plutôt qu'à identifier les apprentissages qu'ils ont fait.		
6 -	Si je le pouvais, je ne donnerais jamais d'examens.		
7 -	Il ne faut pas donner de notes pour un examen formatif.		
8 -	Les enseignants ne sont pas suffisamment rigoureux à l'égard de l'évaluation et laissent passer des étudiants qui ne le devraient pas.		
9 -	Les étudiants étudient plus avant un examen objectif qu'avant un examen à questions ouvertes.		
10-	*L'évaluation est une partie intégrante de la pédagogie.		

11-	Il faut évaluer souvent pour faire en sorte que les étudiants travaillent de façon régulière.		
12-	Les pratiques d'évaluation au cégep encouragent souvent les apprentissages à court terme plutôt que les apprentissages à long terme.		
13-	*L'évaluation n'est pas une activité d'apprentissage.		
14-	J'évalue le rendement scolaire des étudiants uniquement pour satisfaire les exigences administratives de Collège.		
15-	Tout ce que l'étudiant fait en classe et à l'extérieur de la classe mérite d'être évalué donc d'être noté.		

REMARQUES

ÉNONCÉS DE PRATIQUES

		CLARTÉ	PERTINENCE
	Pour aider les étudiants à se préparer à répondre aux questions d'examen, je...		
1 -	fais faire des exercices (pré-tests) avec des questions équivalentes, en salle de classe. (TSRJ)		
2 -	fais des démonstrations (au tableau ou autrement) sur des énoncés équivalents..(TSRJ)		
3 -	donne aux étudiants, à l'avance, une liste de questions et je leur dis que l'examen sera composé de certaines de ces mêmes questions. (TSRJ)		
4 -	fournis les réponses à des questions équivalentes lors de discussions en classe. (TSRJ)		
5 -	fais rédiger une partie des questions d'examen par les étudiants eux-mêmes. (TSRJ)		
6 -	En moyenne, j'évalue les apprentissages, des étudiants dans ce cours... 1. à toutes les semaines 2. aux 2 semaines 3. aux 3 semaines 4. aux 4 semaines 5. aux 5 semaines		
7 -	Quelle proportion du temps de classe l'évaluation des apprentissages de vos étudiants occupe-t-elle? 1. moins de 5% 2. entre 5 et 10% 3. entre 10 et 15% 4. entre 15 et 20% 5. plus de 20%		
8 -	Je donne des points pour les travaux, rapports et/ou exercices que l'étudiant fait à l'extérieur de la classe. (TSRJ)		
9 -	Je donne des points pour les travaux, rapports et/ou exercices que l'étudiant fait en classe. (TSRJ)		

REMARQUES

CATÉGORIE 2: Les cibles de l'évaluation

Il est question ici des objets sur lesquels portent l'évaluation. Les cibles généralement évaluées par les enseignants sont: la connaissance de la matière, le savoir-faire, les habiletés intellectuelles supérieures et les attitudes. Pour Stiggins, l'enseignant devrait savoir comment évaluer chacune de ces cibles.

ÉNONCÉS DE CROYANCES

		CLARTÉ	PERTINENCE
1 -	Il est pratiquement impossible de faire en sorte que les résultats de l'évaluation reflètent les apprentissages des étudiants.		
	Dans la détermination de la note finale quelle importance relative donnez-vous à chacun des items suivants? Indiquez 1 pour le plus important et 5 pour le moins important.		
2 -	Les attitudes (savoir-être)_____		
3 -	Les habiletés, les procédures (savoir-faire)_____		
4 -	Les connaissances (savoir)_____		
5 -	La pensée critique_____		
6 -	La capacité d'analyse et de synthèse et la résolution de problème_____		
7 -	Évaluer, c'est donner un examen sur tout ce qui est dit, tout ce qui est fait en classe.		
8 -	Il n'est pas équitable de poser des questions au-delà de la matière enseignée.		
9 -	Il n'est pas nécessaire que l'évaluation porte sur toute la matière enseignée.		
10-	*Il est impossible, au collégial, d'évaluer autre chose que la connaissance de la matière.		
11-	Il est plus important d'évaluer la compréhension de la matière que la connaissance des faits		
12-	*Au collégial, on devrait surtout évaluer les habiletés intellectuelles supérieures (analyse, synthèse, résolution de problème)		

REMARQUES

ÉNONCÉS DE PRATIQUES

		CLARTÉ	PERTINENCE
1 -	Mes questions d'examen vérifient l'atteinte des objectifs essentiels du cours. (TSRJ)		
2 -	Mes questions d'examen vérifient une partie des objectifs du cours. (TSRJ)		
3 -	Mes questions d'examen vérifient davantage (cochez une réponse) 1. la connaissance d'un contenu disciplinaire. 2. la compréhension de la matière. 3. la capacité d'analyse et de synthèse de l'étudiant. 4. la pensée critique. 5. autres (spécifiez)_____ 6. ne s'applique pas		
	En plus de mesurer et noter l'atteinte des objectifs académiques du cours, je donne des points pour		
4 -	la motivation et l'effort dans l'apprentissage. (TSRJ)		
5 -	le comportement de l'étudiant en salle de classe. (TSRJ)		
6 -	la présence assidue aux cours. (TSRJ)		
7 -	l'exécution des devoirs ou des travaux à la maison ou en classe. (TSRJ)		
8 -	la réalisation de travaux supplémentaires. (TSRJ)		
9 -	les progrès réalisés au cours de la session. (TSRJ)		
10 -	la participation active de l'étudiant en classe. (TSRJ)		

11-	Le choix du <u>contenu</u> des évaluations (quelles habiletés ou connaissances sont mesurées) est fait par: 1. le professeur seul. 2. le professeur et ses collègues donnant le même cours. 3. le professeur et le département. 4. le professeur et les étudiants. 5. le département seul. 6. autres (spécifiez)_____		

REMARQUES

CATÉGORIE 3: Les qualités de l'évaluation

Les caractéristiques d'une "bonne" évaluation varient selon le contexte de l'évaluation. Cependant, quelques critères de qualité sont communs à toutes les situations d'évaluation: la correspondance entre l'objet à évaluer et le dispositif de mesure choisi; le contrôle des sources d'erreur de mesure; la concordance entre l'apprentissage visé et les résultats d'évaluation; l'obtention d'une information dont l'interprétation est claire pour l'enseignant et les étudiants.

Les qualités de l'évaluation ne s'expriment donc pas seulement en termes de fidélité et de validité des instruments de mesure. Les enseignants doivent résoudre des problèmes relatifs à la conception, à l'administration et à la notation des dispositifs de mesure ainsi qu'à l'interprétation des notes et à l'utilisation qui en sera faite.

ÉNONCÉS DE CROYANCES

		CLARTÉ	PERTINENCE
1 -	Après un examen, il est utile de faire l'analyse de mes questions pour en évaluer la qualité.		
2 -	Les questions d'examen sont meilleures lorsqu'elles ont été lues par un second spécialiste de la matière.		
3 -	Tous les étudiants devraient être évalués à partir des mêmes critères.		
4 -	Il est pratiquement impossible de faire en sorte que les résultats d'une évaluation représentent bien les apprentissages des étudiants.		
5 -	Il faut évaluer souvent pour pouvoir se fier davantage à l'ensemble des résultats obtenus.		
6 -	*Mes examens sont efficaces et me disent ce que je veux savoir.		
7 -	*Lorsque j'évalue l'atteinte des objectifs de mon cours, je me sens sûr-e de moi.		
8 -	Il est parfois nécessaire de modifier ses critères d'évaluation au moment de la correction.		
9 -	*Il est nécessaire d'établir ses critères d'évaluation avant la correction.		
10 -	*Il est impossible d'établir ses critères d'évaluation avant le début de la correction.		

	Des critères différents devraient être utilisés pour des groupes différents d'étudiants:		
11-	les plus doués devraient travailler plus fort pour se mériter une note supérieure.		
12-	on devrait être moins exigeant envers les moins doués pour qu'ils aient accès à des notes élevées.		
13-	les étudiants qui exploitent toutes leurs aptitudes devraient avoir des notes plus élevées que ceux qui ne le font pas.		

REMARQUES

ÉNONCÉS DE PRATIQUES

		CLARTÉ	PERTINENCE
1 -	Je fais lire mes questions par un ou des collègues avant d'administrer un examen. (TSRJ)		
2 -	Après la correction d'un examen, je fais l'analyse de mes questions d'examen (clarté, compréhension, difficulté, plausibilité des choix de réponse, etc.). (TSRJ)		
3 -	Lorsque je corrige un examen ou un travail, je me base sur des critères d'évaluation déterminés... 1. à l'avance. 2. au fur et à mesure 3. après une première lecture des copies des étudiants 4. au moment de la correction		
	Dans ce cours, quels facteurs peuvent rendre difficile l'évaluation du rendement scolaire de l'étudiant? (cochez une ou plusieurs réponses)		
4 -	rien		
5 -	la subjectivité du jugement		
6 -	une grande hétérogénéité d'habiletés parmi les étudiants		
7 -	le manque de prérequis chez certains étudiants		
8 -	la difficulté d'évaluer ce qui n'est pas du domaine cognitif		
9 -	autres (spécifiez)_____		
10 -	Lorsque je rédige des questions pour un instrument de mesure, je le fais à partir de questions déjà existantes. (TSRJ)		
11 -	Lorsque je rédige un instrument de mesure, je commence par me faire un plan détaillé de ce que je veux évaluer. (TSRJ)		
12 -	Lorsque je donne un examen, je préviens les étudiants du contenu de l'examen quelques jours à l'avance. (TSRJ)		

13-	Lorsque j'administre un instrument de mesure, je donne les modalités de correction aux étudiants à l'avance. (TSRJ)		
14-	Lorsqu'un étudiant échoue un examen ou un travail, je lui permet de le reprendre. (TSRJ)		
15-	Pour compenser la pratique du «guessing», lors des examens objectifs (à choix de réponses), j'utilise un système de pénalité en enlevant des points dans certaines conditions. (TSRJ)		
16-	Combien de temps prenez-vous en moyenne par semaine pour la PRÉPARATION de tout ce qui fait partie de l'évaluation du rendement scolaire de l'étudiant de ce cours? (cochez une réponse) 1. pas plus de 3 heures par semaine 2. entre 3 heures et 6 heures 3. entre 6 heures et 9 heures 4. entre 9 heures et 12 heures 5. plus de 12 heures		

REMARQUES

CATÉGORIE 4: Les instruments d'évaluation

Les professeurs utilisent au moins trois formes d'évaluation des apprentissages en salle de classe:.les instruments de type "papier-crayon", l'observation et l'échange verbal. Les professeurs doivent savoir se servir de ces instruments tout en comprenant que les règles de validité sont différentes d'une forme à l'autre.

ÉNONCÉS DE CROYANCES

		CLARTÉ	PERTINENCE
1 -	*Seules les questions à développement peuvent mesurer les apprentissages importants.		
2 -	Tous les examens devraient être à livre ouvert.		
3 -	Toutes les évaluations devraient être des autoévaluations.		
4 -	Les questions à choix multiples privilégient l'évaluation des connaissances mémorisées.		
5 -	Des examens-synthèses de fin de session devraient être obligatoires dans presque toutes les disciplines.		
6 -	Les questions à choix multiples peuvent mesurer la compréhension de la matière.		
7 -	Il est presqu'impossible de rédiger des questions d'examen qui mesurent les habiletés intellectuelles supérieures.		
8 -	Seuls les travaux de recherche ou la réalisation d'un projet mesurent vraiment l'atteinte des objectifs.		
9 -	*Les examens avec questions à choix multiples mesurent mieux les apprentissages importants que les questions à développement.		

REMARQUES

ÉNONCÉS DE PRATIQUES

		CLARTÉ	PERTINENCE
	Que ce soit pour des fins formative ou sommative, utilisez-vous les approches suivantes? (OUI/NON)		
1 -	devoirs		
2 -	dossier anecdotique, journal de bord		
3 -	examen surprise		
4 -	observation directe avec grille et échelle d'appréciation		
5 -	jeux questionnaires		
6 -	rapport d'activités/laboratoire		
7 -	travail de recherche		
8 -	questions orales		
9 -	liste de vérification		
10 -	projet thématique		
11 -	problèmes à résoudre		
12 -	production ou démonstration		
13 -	travail en classe		
14 -	phrases à compléter		
15 -	questions à appariement, association		
16 -	questions à choix multiples		
17 -	questions demandant un développement court		
18 -	questions demandant un développement long		
19 -	l'auto-évaluation		
20 -	l'évaluation par les pairs		
21 -	autres(spécifiez)_____		

22- C 4	Laquelle de ces approches vous permet le mieux de savoir si vos étudiants, à la fin de la session, ont atteint les objectifs visés? _____		
23- C 4	Laquelle vous informe le mieux des progrès de l'étudiant? _____		
24- C 4	Laquelle de ces approches utiliseriez-vous davantage si vous la connaissiez mieux? _____		
26- C 4	Laquelle de ces approches utiliseriez-vous davantage si vous aviez plus de temps? _____		
27- C 4	Laquelle de ces approches vous semble la plus utile pour aider vos étudiants à apprendre? _____		
28- C 4	Laquelle de ces approches semble la plus utile pour accumuler des notes en vue du bulletin de fin de session?_____		
30-	Dans mon département, le choix des types d'instruments est fait par: le professeur seul. le professeur et ses collègues donnant le même cours. le professeur et le département. le professeur et les étudiants. le département seul. Autres (spécifiez)_____		
  31- 32- 33- 34-	Pendant un examen, je présente mes questions d'examen de la façon suivante: j'écris les questions au tableau. (TSRJ) je les projette sur acétate. (TSRJ) je les imprime et en distribue un exemplaire à chaque étudiant. (TSRJ) j'en fais la lecture à haute voix et les étudiants les prennent en note. (TSRJ)		

	Dans mes examens à développement,		
35-	je propose des questions «au choix» (ex: «répondre à 3 questions parmi un choix de 8»). (TSRJ)		
36-	je permet aux étudiants de répondre «à la maison» et de remettre leurs réponses lors de la prochaine rencontre. (TSRJ)		
37-	je permet aux étudiants d'avoir en main toute documentation utile. (TSRJ)		
38-	autres (spécifiez)_____		

REMARQUES

CATÉGORIE 5: Les dimensions interpersonnelles de l'évaluation

Les variables interpersonnelles ont un impact sur l'évaluation en salle de classe et sur l'interprétabilité des notes. Certaines contraintes interpersonnelles viennent des étudiants eux-mêmes, d'autres des enseignants. (attentes et attitudes à l'égard des étudiants; conceptions de leur rôle dans l'enseignement, l'apprentissage, l'évaluation et le contrôle de la salle de classe; le tempérament; la motivation à enseigner et les aptitudes à réagir aux personnes). L'enseignant compétent doit savoir reconnaître et contrôler ces nombreuses dimensions personnelles s'il veut assurer à ses étudiants une évaluation équitable.

ÉNONCÉS DE CROYANCES

		CLARTÉ	PERTINENCE
1 -	*Il faut s'abstenir d'augmenter ou de diminuer la note d'un étudiant pour l'encourager ou le pousser à travailler davantage.		
	*La note finale transmise au bulletin pourrait être augmentée pour récompenser:		
2 -	la participation active de l'étudiant en classe.		
3 -	l'effort fourni.		
4 -	les progrès de l'étudiant sur l'ensemble de la session (l'évolution).		
5 -	la créativité de l'étudiant.		
6 -	la présence à tous les cours.		
7 -	Lors de la correction des questions à développement, la connaissance de l'identité du répondant peut m'influencer.		

	La note finale transmise au bulletin pourrait être diminuée pour pénaliser:		
8 -	des absences non-motivées		
9 -	l'absence ou le peu d'effort fourni par l'étudiant		
10 -	l'indiscipline en classe		
11 -	le plagiat		

REMARQUES

ÉNONCÉS DE PRATIQUES

		CLARTÉ	PERTINENCE
1 -	Il m'arrive de donner un examen facile pour compenser un examen antérieur difficile. (TSRJ)		
2 -	J'ajuste mes exigences en fonction des différents niveaux d'habiletés de mes étudiants. (TSRJ)		
3 -	J'ai tendance à encourager un étudiant faible en accordant une note de passage lorsque ses efforts me semblent appréciables. (TSRJ)		
4 -	Il m'arrive de modifier les notes obtenues par certains étudiants. (TSRJ)		
5 -	Il m'arrive d'ajuster à la hausse ou à la baisse mes moyennes de classe. (TSRJ)		
6 -	Dans mon calcul des notes qui vont au bulletin, j'ajoute à l'évaluation de l'atteinte des objectifs, une note correspondant à l'effort consenti par l'étudiant pendant la session. (TSRJ)		
7 -	Dans l'attribution des notes, j'ai tendance à être plus indulgent envers les étudiants moins doués. (TSRJ)		
8 -	Mes évaluations sont pondérées de façon à donner une moyenne de classe, en fin de sesssion 1. entre 60 et 70% 2. entre 70 et 80% 3. plus de 80% 4. pas inférieure à 60% 5. pas inférieure à 50% 6. je ne modifie jamais la moyenne de classe		

REMARQUES

CATÉGORIE 6: La rétroaction et la notation

Les enseignants fournissent continuellement des rétroactions à partir des résultats des évaluations.

Les pratiques de notation sont plus appropriées lorsque les professeurs: informent à l'avance les étudiants de leur politique et de leur attitude à ce sujet; construisent la note finale autour des variables pertinentes; introduisent dans la note finale des résultats fondés sur des évaluations de qualité; maintiennent des données complètes et appropriées quand aux apprentissages de leurs étudiants; pondèrent correctement les diverses notes accumulées et établissent convenablement les seuils de réussite afin de traduire en note les compétences atteintes.

Quant aux autres formes de rétroaction, les professeurs ont des pratiques appropriées lorsqu'ils s'assurent que leur rétroaction correspond à ce qui est attendu explicitement dans l'objectif visé, est transmise dans le temps approprié et est signifiante pour l'étudiant.

ÉNONCÉS DE CROYANCES

		CLARTÉ	PERTINENCE
1 -	La note moyenne de la classe est un reflet direct de la qualité de l'enseignement.		
2 -	Dans un groupe, la distribution des notes devrait suivre la courbe normale, en forme de cloche: les notes les meilleures et les moins élevées devraient être réservées à quelques-uns seulement.		
3 -	De façon générale, les critères de réussite ou d'échec, dans mon collège, sont: 1. trop indulgents, généreux 2. trop stricts 3. adéquats		
4 -	La notation est un handicap à l'enseignement.		
5 -	Certains professeurs évaluent et notent parce qu'ils en sont obligés et, en conséquence, le font pour s'en débarrasser.		
6 -	*Les notes que j'attribue représentent peu ce que mes élèves ont appris.		
7 -	Il faut tenir compte, dans les notes, des résultats bons ou mauvais obtenus par l'étudiant lors d'une évaluation effectuée au début d'un apprentissage.		
8 -	Il faut éviter de faire des évaluations qui impliquent le jugement personnel, la subjectivité de l'enseignant.		

9 -	Les notes devraient refléter le nombre ou le pourcentage des objectifs maîtrisés par mes étudiants.		

REMARQUES

ÉNONCÉS DE PRATIQUES

		CLARTÉ	PERTINENCE
1 -	La note finale inscrite au bulletin de l'étudiant contient, entre autres, des notes obtenues dans des évaluations formatives. (TSRJ)		
2 -	Dans la détermination des notes finales, je tiens compte de la politique d'évaluation des apprentissages adoptée par mon département ou par mon collège. (TSRJ)		
3 -	Lorsque je retourne un examen ou un travail aux étudiants, je leur donne une rétroaction détaillée sur les points à améliorer. (TSRJ)		
4 -	Après un test, je prends les moyens nécessaires pour que les étudiants puissent améliorer leurs points faibles. (TSRJ)		
5 -	Je retourne les tests et travaux mineurs corrigés aux étudiants après: 1. une semaine 2. deux semaines 3. trois semaines 4. quatre semaines 5. 5 semaines		
6 -	Combien de temps prenez-vous en moyenne par semaine pour la CORRECTION de tout ce qui fait partie de l'évaluation du rendement scolaire de l'étudiant de ce cours? (cochez une réponse) 1. pas plus de 3 heures 2. plus de 3 heures mais pas plus de 6 heures 3. plus de 6 heures mais pas plus de 9 heures 4. plus de 9 heures mais pas plus de 12 heures 5. plus de 12 heures		

REMARQUES

Formation et perfectionnement

		CLARTÉ	PERTINENCE
	Quelle préparation formelle avez-vous reçue en matière d'ÉVALUATION DES APPRENTISSAGES? (Vous pouvez cochez plus d'une réponse)		
1 -	Aucune		
2 -	Au moins un cours au baccalauréat		
3 -	Au moins un cours au CPEC-Performa		
4 -	Au moins un cours à la maîtrise		
5 -	Au moins une journée pédagogique		
6 -	Au moins une conférence ou un atelier		
7 -	Autres (spécifiez)		
	Quelle évaluation faites-vous de vos connaissances dans les domaines suivants?		
8 -	Évaluation formative Excellente/Bonne/Moyenne/Passable/Nulle		
9 -	Évaluation sommative Excellente/Bonne/Moyenne/Passable/Nulle		

	Qu'est-ce qui pourrait vous aider au chapitre de l'évaluation des apprentissages de vos étudiants? (cochez plusieurs réponses s'il y a lieu)		
	des sessions de formation en:		
10-	Spécification d'objectifs		
11-	Construction d'examens		
12-	Techniques d'observation		
13-	Évaluation formative		
14-	Évaluation critériée		
15-	Évaluation des attitudes		
16-	Planification de l'évaluation		
17-	La notation		
18-	La correction (annotation)		
19-	Autres (spécifiez)_____		
	des ressources telles:		
20-	Une banque de questions d'examens		
21-	Des examens de fin de session déjà construits		
22-	Des examens modèles		
23-	Autres (spécifiez)_____		

REMARQUES:

Informations générales

Nom du collège:_____

Quels sont les trois **premiers chiffres des cours** que vous enseignez habituellement:

_ _ _

Sexe: 1. féminin _____ 2. masculin _____

Age:
1. Moins de 30 ans _____
2. 30 à 39 ans _____
3. 40 à 49 ans _____
4. 50 et plus _____

Années d'expérience reconnues dans l'enseignement:

1. Moins de 5 ans _____
2. 5 à 10 ans _____
3. 11 à 20 ans _____
4. Plus de 20 ans _____

Années de scolarité reconnues pour fins de rémunération:

1. Moins de 16 _____
2. 16 ans _____
3. 17 ans _____
4. 18 ans _____
5. 19 ans _____

COLLÈGE
MONTMORENCY

475, boulevard de l'Avenir
Laval, Québec
H7N 5H9

Téléphone:
(514) 667-5100

Télécopieur:
(514) 336-0835

Laval, le 31 janvier 1992

Dans le cadre de son programme PAREA de subvention à la recherche pédagogique, le Service de la recherche de la DGEC a accepté de subventionner notre projet de recherche intitulé:

> ***«Étude des croyances et des pratiques des enseignants de cégeps à l'égard de l'évaluation des apprentissages.»***

Nous préparons un questionnaire qui sera administré, cette session, à un échantillon de professeurs de l'ensemble du réseau des cégeps. Nous arrivons aujourd'hui à une étape-clé de notre processus. Nous voulons valider notre questionnaire, avant de l'imprimer et de l'envoyer à près de mille professeurs. C'est pour cette raison que nous vous écrivons et sollicitons votre collaboration. D'abord, permettez-nous de vous décrire le contexte général.

Le projet:

Notre étude descriptive, d'une durée de deux ans, vise l'atteinte des objectifs suivants:

- Identifier les croyances et les pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages et ce, en situation régulière d'enseignement et d'apprentissage, en salle de classe.

- Explorer l'existence de liens entre les croyances et les pratiques.

Les résultats de cette étude que nous pensons publier en juin 1993, permettront de mieux comprendre les différents paramètres de l'évaluation des apprentissages au collégial et, notamment, de guider la planification d'activités de perfectionnement adaptées aux besoins des enseignants et enseignantes des cégeps.

La méthodologie

Les données nécessaires à cette recherche seront principalement recueillies par questionnaire. L'automne dernier, une revue de la littérature ainsi que des entrevues auprès de douze (12) professeurs de cégeps nous ont permi de dégager un certain nombre d'énoncés de questions, portant tantôt sur les croyances à propos de l'évaluation, tantôt sur les pratiques en salle de classe.

Ces énoncés ont été soumis, en janvier, à un groupe d'experts qui en ont jugé la clarté et la pertinence. Ceci nous amène à rédiger notre questionnaire dans sa forme finale et à l'administrer, à l'essai, à un petit groupe de professeurs afin de roder l'ensemble des dimensions critiques propres à la cueillette de données par questionnaire: durée de la tâche de réponse, lisibilité, clarté, convivialité, saisie des données, compréhension des consignes, etc.. C'est précisément pour cette étape de validation finale que nous vous sollicitons aujourd'hui.

C'est après avoir corrigé les dernières lacunes, que le questionnaire sera distribué à un millier de professseurs dans une quinzaine de cégeps représentatifs de l'ensemble des collèges et de l'ensemble des disciplines enseignées au secteur régulier. Toutes ces personnes auront été identifiées par échantillonnage aléatoire stratifié.

La tâche

Si vous acceptiez de faire partie de la vingtaine de professeurs qui nous aiderons à roder notre questionnaire, nous vous transmettrons l'ensemble des documents suivant: une lettre d'introduction et de présentation du projet, ainsi que le questionnaire lui-même, dans sa forme "finale", avec consignes pour y répondre. À celà s'ajoutera une feuille vous permettant d'indiquer le temps que vous aurez mis à répondre au questionnaire, vos commentaires, vos suggestions, etc.. Nous visons vous faire parvenir le questionnaire à la fin de février.

Vos réponses seront bien sûr complétées et analysées. Ceci nous permettra de roder tout le processus de saisie des données et de traitement statistique sur ordinateur. Il va de soi que, à l'instar des réponses que nous fourniront l'ensemble des répondants, toutes vos réponses au questionnaire ainsi que vos commentaires-feedback seront strictement confidentiels.

Puisque le questionnaire comportera près de 80 questions, nous espérons que le tout ne vous prendra pas plus de deux heures de votre temps.

De même que nous communiquons avec chacun d'entre vous séparément, chacun répondra au questionnaire individuellement et nous vous laisserons un délai d'une dizaine de jours pour nous retourner vos réponses.

Nous savons que vous êtes très sollicité-e, notamment par votre tâche d'enseignement ...et d'évaluation, et nous comprenons fort bien que tout ceci peut arriver à un moment inconvenant. C'est pourquoi nous tenons à ce que vous soyez bien à l'aise d'accepter ou de refuser de participer à cette validation de questionnaire.

Nous vous serions reconnaissants de communiquer avec nous (poste 189) ou avec la secrétaire, Nicole Brunet (poste 192), pour nous faire savoir si nous pouvons compter sur votre participation à la fin de février, probablement dans la semaine du 24 février.

Dans l'attente de votre réponse, veuillez agréer nos salutations.

Louise Ménard,
Professeure, département
de Soins infirmiers

Robert Howe,
Conseiller pédagogique

ANNEXE 3

LE PROTOCOLE D'ENTREVUE

Les entrevues de pré-consultation

Profil des professeurs interviewés, en pré-consultation, à l'automne 1991

Les interviews ont été menées auprès d'un professeur de chacune des disciplines suivantes:

Secteur général:

- langues modernes
- mathématiques
- physique
- français
- psychologie
- politique

Secteur professionnel:

- soins infirmiers
- électrotechnique
- techniques administratives
- service d'éducation en service de garde (cégep voisin)
- techniques d'architecture
- technique de gestion de bureau

Le protocole d'entrevues individuelles, automne 1991

Liste des questions qui ont servi à guider les entrevues.

I- Le "Pourquoi"

1. Est-ce utile d'évaluer? Quelles sont les raisons d'être de l'évaluation des apprentissages au cégep?

2. Selon vous, les évaluations que vous faites ont-elles des impacts favorables ou défavorables sur vos étudiants? Quels sont les effets de l'évaluation sur les étudiants? Est-ce que l'évaluation aide ou nuit à l'apprentissage des étudiants?

3. Trouvez-vous nécessaire de faire de l'évaluation formative au collégial?

II- Le "Quoi"

1. Qu'est-ce que vous évaluez chez vos étudiants?

2. Favorisez-vous l'usage d'un examen-synthèse à la fin de la session?

3. Quelle est votre politique à l'égard de l'effort, de la participation en classe, du comportement en classe, des progrès réalisés pendant la session?

4. Qu'est-ce qui entre dans votre note finale?

III- Le "Comment"

1. Comment préparez-vous vos examens? Comment ajustez-vous la durée de l'examen avec le contenu de l'examen?

2. Comment évaluez-vous les apprentissages? Quels outils? Quelles stratégies?

3. Pourquoi choisissez-vous ces outils en particuliers?

4. Que pensez-vous des questions à choix multiples? Des questions à développement? Pourquoi? Qu'est-ce qu'on peut évaluer avec ces types de questions?

5. Tenez-vous compte du stress des étudiants, de la fatigue?

6. Cherchez-vous à ajuster vos moyennes de groupes en faisant volontairement des examens plus difficiles ou plus faciles? Pourquoi?

7. Évaluez-vous les étudiants à partir de critères pré-établis?

8. Vous arrive-t-il de modifier votre barême de correction pour atteindre une moyenne de groupe qui vous convient?

9. Où avez-vous appris ce que vous savez en mesure et évaluation?

IV- Les autres

Connaissez-vous des croyances et des pratiques d'évaluation différentes des vôtres? Quelles sont-elles?

V- Le perfectionnement

Si on organisait du perfectionnement en matière d'évaluation des apprentissages, quels seraient, selon vous, les sujets prioritaires à traiter, les problèmes les plus urgents?

ANNEXE 4
POPULATION ET ÉCHANTILLON

STRATES

11 strates d'origine			Strates finales		
CODES	Appellations d'origine	Pondérations	CODES	Appellations complètes	Appellations abrégées
101	Sciences biologiques	(0.807963)	101	Sciences biologiques	Sc. biologiques
			109	Éducation physique	Éduc. physique
199	Techniques biologiques	(1.136783)	199	Techniques biologiques	Tech. biologiques
201	Sciences physiques	(0.993447)	201	Sciences physiques	Sc. physiques
299	Techniques physiques	(1.007684)	299	Techniques physiques	Tech. physiques
301	Sciences humaines (incluant Philosophie)	(2.053991)	301	Sciences humaines (incluant Philosophie)	Sc. humaines
399	Techniques humaines	(0.716172)	399	Techniques humaines	Tech. humaines
401	Administration (pré-universitaire)	(0.639299)	400	Administration (pré-universitaire et technique)	Administration
499	Administration (technique)	(0.666154)			
501	Arts (pré-universitaire)	(1.359259)	500	Arts (pré-universitaire et technique)	Arts
599	Arts (technique)	(0.607543)			
601	Lettres	(1.074919)	601	Lettres	Lettres

Note a: La ligne "TOTAL" au bas de chaque tableau stratifié sera appelée "**Échantillon total**".

Note b: La codification originelle des strates prévoyait que, pour chacun des six grands champs de connaissance (premier chiffre), le suffixe 01 devait désigner les disciplines du secteur général et le suffixe 99, les disciplines du secteur professionnel. Dans le premier champ, la discipline Éducation physique a été isolée pour des motifs de spécificité pédagogique, tel qu'expliqué au chapitre 4. De plus, les disciplines des champs 4 et 5 ont été regroupées, d'où le suffixe unique 00.

Les cégeps participants

Collèges	Échantillon prévu	Taux erreur SPOC	Échantillon final	Retours, au 6 juin 1992	Taux de réponses final
ANDRÉ-LAURENDEAU	29	,1	27	12	,44
ABITIBI-TÉMISCAMINGUE	32	----	32	17	,53
AHUNTSIC	82	----	82	22	,27
ALMA	22	,05	22	10	,45
RÉGION DE L'AMIANTE	19	----	19	9	,47
BOIS-DE-BOULOGNE	25	,08	23	9	,39
Champlain-LENNOXVILLE	13	----	13	1	,08
Champlain-ST-LAMBERT	24	,21	22	3	,14
Champlain-ST-LAURENCE	9	----	9	4	,44
CHICOUTIMI	34	,06	32	18	,56
DAWSON	69	----	69	15	,22
DRUMMONDVILLE	37	,22	37	13	,35
ÉDOUARD-MONTPETIT	66	----	66	26	,39
FRANÇOIS-XAVIER-GARNEAU	54	----	54	26	,48
GASPÉ	32	,19	32	20	,62
GRANBY	13	----	13	8	,62
HAUTERIVE	15	----	15	3	,2
HÉRITAGE	13	----	13	1	,08
JOHN ABBOTT	54	----	54	8	,15
JOLIETTE-DE-LANAUDIÈRE	35	,03	34	15	,44
JONQUIÈRE	54	,06	51	26	,51
LIONEL-GROULX	39	----	39	8	,21
LA POCATIÈRE	20	----	20	13	,65
LÉVIS	38	----	38	16	,42
LIMOILOU	49	----	49	21	,43

Les cégeps participants (suite)

Collèges	Échantillon prévu	Taux erreur SPOC	Échantillon final	Retours, au 6 juin 1992	Taux de réponses final
MAISONNEUVE	55	,09	53	14	,26
MATANE	17	----	17	7	,41
MONTMORENCY	46	----	46	23	,50
OUTAOUAIS	44	,05	42	24	,57
RIMOUSKI	34	,12	30	20	,67
RIVIÈRE-DU-LOUP	23	,09	21	11	,52
ROSEMONT	22	----	22	9	,41
ST-JÉRÔME	37	----	37	21	,57
SEPT-ILES	13	----	13	6	,46
SHAWINIGAN	20	,15	18	11	,61
SHERBROOKE	65	,09	60	25	,42
SOREL-TRACY	20	,1	19	11	,58
ST-HYACINTHE	23	,13	23	10	,43
ST-FÉLICIEN	8	----	8	4	,5
ST-JEAN-SUR-RICHELIEU	28	,07	26	17	,65
ST-LAURENT	32	----	32	13	,41
STE-FOY	84	,14	76	36	,47
TROIS-RIVIÈRES	58	----	58	31	,53
VIEUX MONTRÉAL	85	----	85	23	,27
VALLEYFIELD	17	----	17	9	,53
VANIER	66	----	66	10	,15
VICTORIAVILLE	16	----	16	8	,5
Totaux	**1690**	**,106 ***	**1510 ****	670	44,3

Notes: * Taux d'erreur moyen, pour 19 collèges qui ont déclaré des erreurs dans l'échantillon prévu.
** Nombre estimé de professeurs rejoints si on généralise le taux d'erreur SPOC à tous les collèges.

ANNEXE 5

TABLEAUX D'ANALYSES STRATIFIÉES

Présentation

La présentation des données, au chapitre 5, contient les résultats obtenus pour l'ensemble de l'échantillon. Mais lorsque l'analyse de variance (ANOVA) nous révèle qu'il y a différence significative de réponses entre les strates, nous présentons une ventilation des réponses obtenues pour chaque strate ainsi que les réponses de l'échantillon total, pour fins de comparaison.

Tous ces tableaux sont issus de tableaux de contingence entre les dix strates et chaque variable concernée.

Pour des raisons de lisibilité, les énoncés de questions sont abrégés et le lecteur est invité à se reporter au questionnaire, en annexe 1, pour y trouver l'énoncé complet, tel que présenté aux répondants.

Pour les mêmes raisons de lisibilité, les échelles de mesure sont abrégées selon la légende suivante:

a) Échelle de fréquence:

	Toujours	Souvent	Parfois	Rarement	Jamais
abrégé:	T	S	P	R	J

b) Échelle d'accord ou de désaccord:

	Très en accord	En accord	En désaccord	Très en désaccord	Ne sais pas
abrégé:	TA	A	D	TD	NSP

Dans les strates, les proportions extrêmes, habituellement des maxima, sont mises en évidence afin d'illustrer les écarts par rapport à la réponse moyenne de l'échantillon total.

Tous les tableaux qui suivent ont une numérotation qui renvoie, dans le chapitre 5, à la description de croyances "C" ou de pratiques "P" et ce, dans chacune des six catégories de Stiggins. Ainsi, le tableau C2-b analyse des croyances de catégorie 2. Le tableau P5 analyse des pratiques de catégorie 5. Il en va ainsi de suite.

Les tableaux stratifiés de cette annexe portent le suffixe "b" par opposition aux tableaux du chapitre 5 qui, eux, portent le suffixe "a". Ainsi, au chapitre 5, le premier tableau descriptif est numéroté C1-a alors que dans la présente annexe, le tableau stratifié de même catégorie est numéroté C1-b. Il en va de même pour les tableaux sur les pratiques: P6-a au chapitre 5 et P6-b à l'annexe 5.

Les calculs ne contiennent pas les "ne sais pas" et sont basés sur les fréquences brutes. C'est ce qui peut expliquer certaines différences entre ces tableaux d'analyse par strates et les tableaux présentant les fréquences pondérées des répondants dans leur ensemble. Se référer au chapitre 5, tableau C5-a, pour y trouver les proportions de "ne sais pas" ainsi que les proportions de l'échantillon total.

Tableau C1-b: **Croyances relatives à l'utilisation de l'évaluation en classe: Analyse des questions 6d) et 6j)**

	Q6d) Suite à l'évaluation, les enseignants devraient être disposés à réajuster leur contenu de cours N = 616 p = ,0001				Q6j) Il n'est pas nécessaire de faire de l'évaluation formative N = 608 p = ,0004		
	TA	A	D	TD	A	D	TD
Sc. biologiques	18%	59%	18%	6%	6%	53%	41%
Éduc. physique	52%	42%	6%	----	3%	32%	65%
Tech. biologiques	41%	50%	7%	3%	1%	20%	79%
Sc. physiques	12%	29%	40%	19%	7%	47%	46%
Tech. physiques	25%	52%	17%	5%	2%	45%	53%
Sc. humaines	20%	65%	14%	2%	4%	58%	38%
Tech. humaines	55%	40%	6%	----	25	17%	81%
Administration	39%	42%	18%	1%	4%	48%	47%
Arts	40%	45%	12%	3%	2%	45%	53%
Lettres	44%	41%	11%	3%	5%	48%	48%
Échantillon total:	**34%**	**45%**	**16%**	**5%**	**4%**	**41%**	**55%**
					Note: Les colonnes TA et A sont combinées		

Note: Les calculs ne contiennent pas les "ne sais pas" et sont basés sur les fréquences brutes. C'est ce qui peut expliquer certaines différences entre ces tableaux d'analyse par strates et les tableaux présentant les fréquences pondérées des répondants dans leur ensemble. Se référer au chapitre 5, tableau C1-a, pour y trouver les proportions de "ne sais pas" ainsi que les proportions de l'échantillon total.

Tableau C2-b: Croyances relatives aux cibles de l'évaluation

	Q6l) Examens devraient porter sur tout ce qui est dit en classe N = 617; p = ,0129			Q6o) Est impossible d'évaluer autre chose que la connaissance N = 606; p = ,016				Q6p) Pratiquement impossible d'évaluer les attitudes N = 557; p = ,0001				Q6r) Devrait surtout évaluer les habiletés intellectuelles supérieures N = 594; p = ,0475		
	A	D	TD	TA	A	D	TD	TA	A	D	TD	TA	A	D
Sc. biologiques	38%	53%	9%	----	3%	62%	34%	7%	17%	57%	20%	9%	**55%**	36%
Éduc. physique	27%	53%	20%	----	3%	53%	43%	----	16%	45%	**39%**	3%	41%	55%
Tech. biologiques	21%	**57%**	**22%**	3%	1%	53%	43%	3%	7%	59%	32%	7%	47%	46%
Sc. physiques	48%	49%	3%	----	10%	62%	28%	5%	**37%**	52%	6%	6%	49%	45%
Tech. physiques	43%	40%	17%	3%	**15%**	58%	24%	4%	25%	**62%**	9%	3%	33%	**63%**
Sc. humaines	28%	60%	11%	4%	**13%**	57%	26%	**14%**	30%	43%	14%	**17%**	40%	44%
Tech. humaines	37%	47%	16%	----	2%	42%	**57%**	2%	12%	48%	38%	4%	35%	**61%**
Administration	42%	48%	10%	1%	8%	59%	32%	3%	25%	55%	17%	9%	41%	50%
Arts	46%	43%	11%	2%	4%	**67%**	28%	3%	21%	48%	28%	7%	45%	48%
Lettres	29%	60%	11%	3%	12%	53%	31%	11%	22%	55%	13%	**17%**	51%	32%
Échantillon total:	**37%**	**51%**	**13%**	**2%**	**8%**	**57%**	**34%**	**5%**	**21%**	**53%**	**21%**	**8%**	**44%**	**48%**
	Note: Les colonnes A et TA sont ici combinées											Note: Les colonnes D et TD sont ici combinées		

Note: Les calculs ne contiennent pas les "ne sais pas" et sont basés sur les fréquences brutes. C'est ce qui peut expliquer certaines différences entre ces tableaux d'analyse par strates et les tableaux présentant les fréquences pondérées des répondants dans leur ensemble. Se référer au chapitre 5, tableau C2-a, pour y trouver les proportions de "ne sais pas" ainsi que les proportions de l'échantillon total.

Tableau P2-b: **Analyse, strate par strate, des habiletés mesurées, selon la question 2: "Mes instruments d'évaluation visent la mesure de:" (se référer à l'annexe 4 ou au chapitre 4 pour y voir les appellations des 10 strates)**

Q2a)	La connaissance d'un contenu disciplinaire N = 608; p = ,0036	Échantillon total	101	109	199	201	299	300	399	400	500	600
	toujours	47%	53%	50%	39%	**61%**	33%	55%	28%	49%	47%	52%
	souvent	37%	38%	33%	**45%**	33%	**46%**	28%	**54%**	38%	28%	29%
	parfois	11%	6%	17%	11%	5%	**18%**	8%	14%	12%	11%	14%
	rarement ou jamais	5%	3%	----	5%	1%	3%	9%	4%	2%	**14%**	6%
Q2c)	L'application des procédures (savoir-faire) N = 619; p = ,0001											
	toujours	37%	29%	**61%**	26%	35%	29%	13%	35%	41%	**56%**	52%
	souvent	41%	21%	35%	51%	50%	**54%**	35%	40%	44%	29%	33%
	parfois	18%	**35%**	3%	18%	14%	14%	**43%**	23%	14%	14%	12%
	rarement ou jamais	3%	**15%**	----	5%	1%	3%	9%	2%	----	2%	3%
Q2d)	Résolution de problèmes complexes N = 617; p = ,0001											
	toujours	8%	3%	3%	3%	9%	2%	4%	6%	9%	**25%**	15%
	souvent	30%	**49%**	23%	29%	38%	30%	27%	24%	34%	30%	21%
	parfois	38%	34%	29%	43%	42%	46%	**51%**	36%	42%	25%	26%
	rarement	18%	14%	29%	16%	10%	20%	18%	**30%**	14%	12%	24%
	jamais	5%	----	**16%**	10%	1%	3%	----	4%	1%	7%	14%

Tableau P2-b: (suite) Q2: Mes instruments d'évaluation visent la mesure de:

Q2e) Analyse et synthèse N = 616; p = ,0001	Échantillon total	101	109	199	201	299	300	399	400	500	600
toujours	19%	14%	10%	8%	17%	11%	28%	8%	15%	**33%**	41%
souvent	46%	**54%**	19%	43%	47%	53%	46%	48%	51%	36%	44%
parfois	27%	20%	**45%**	42%	30%	29%	20%	40%	21%	22%	9%
rarement ou jamais	9%	11%	**26%**	8%	7%	6%	6%	4%	12%	9%	6%
Q2f) La pensée critique N = 614; p = ,0001											
toujours	10%	3%	10%	8%	5%	5%	20%	2%	6%	**27%**	18%
souvent	30%	20%	17%	29%	20%	26%	35%	40%	29%	36%	**42%**
parfois	37%	37%	**47%**	41%	40%	44%	31%	40%	41%	18%	25%
rarement	18%	**29%**	17%	14%	26%	19%	11%	13%	19%	15%	11%
jamais	6%	**11%**	**10%**	8%	9%	6%	2%	4%	5%	4%	5%
Q2g) Les attitudes (savoir-être) N = 611; p = ,0001											
toujours	11%	6%	**48%**	16%	----	5%	4%	17%	5%	25%	6%
souvent	24%	11%	35%	**42%**	11%	27%	19%	37%	14%	36%	18%
parfois	26%	20%	13%	27%	25%	20%	24%	**33%**	31%	23%	32%
rarement	23%	**43%**	3%	9%	31%	28%	30%	12%	33%	9%	22%
jamais	17%	20%	----	6%	**33%**	20%	24%	2%	17%	7%	22%

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total. Les échelles "rarement" et "jamais" ont été combinées pour certains items, afin de valider les tableaux.

Tableau P2-b: **(suite) Analyse stratifiée des réponses aux questions 1 et 3**

	Énoncé	N	Échantillon total	101	109	199	201	299	300	399	400	500	600
Q1)	Je mesure... p = ,0101	604											
	tout le contenu		51%	44%	32%	44%	68%	35%	55%	50%	56%	60%	46%
	un échantillon		23%	35%	35%	25%	18%	23%	19%	21%	16%	21%	28%
	seulement l'essentiel		26%	21%	32%	31%	14%	42%	26%	29%	28%	19%	26%

Q3) **Dans ce cours, en plus d'attribuer des points pour les objectifs liés à la matière enseignée, j'en accorde pour:**

Rang	Énoncé	N	Échantillon total	101	109	199	201	299	300	399	400	500	600
1	g- le respect de mes exigences de présentation des travaux	617	73%	66%	57%	62%	70%	88%	60%	85%	81%	88%	59%
2	c- l'exécution de devoirs ou d'exercices à la maison ou en classe	621	70%	63%	55%	56%	56%	69%	72%	75%	83%	85%	76%
3	f- la créativité, l'originalité	615	53%	43%	37%	39%	33%	57%	48%	74%	53%	83%	71%
4	a- la motivation et l'effort dans l'apprentissage	611	48%	20%	94%	39%	29%	64%	43%	48%	40%	78%	48%
5	e- la participation active de l'étudiant en classe	616	40%	14%	97%	34%	13%	39%	33%	61%	32%	71%	47%
6	b- la présence assidue aux cours	616	37%	11%	87%	42%	14%	36%	37%	49%	26%	65%	33%
7	d- les progrès réalisés au cours de la session (évolution)	612	36%	14%	77%	33%	15%	38%	32%	30%	30%	69%	48%

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total (p = ,0001).

Tableau C3-b: La validation des items

	Q33a) Analyse des questions N = 621 p = ,0001			Q33b) Questions d'examen lues par un second spécialiste N = 621 p = ,0001				Q33c) La clarté des consignes des travaux écrits N = 621 p = ,0008			
	TA	A	D	TA	A	D	NSP	TA	A	D	NSP
Sc. biologiques	38%	56%	6%	9%	65%	21%	6%	12%	74%	6%	9%
Éduc. physique	40%	60%	----	27%	67%	3%	3%	27%	67%	7%	----
Tech. biologiques	**61%**	32%	3%	**36%**	52%	5%	6%	**41%**	55%	3%	1%
Sc. physiques	31%	60%	7%	10%	57%	**24%**	8%	20%	62%	11%	7%
Tech. physiques	27%	67%	5%	17%	58%	19%	6%	14%	65%	**17%**	3%
Sc. humaines	27%	62%	5%	5%	64%	15%	**16%**	7%	67%	15%	**11%**
Tech. humaines	48%	48%	4%	11%	63%	13%	13%	19%	65%	11%	6%
Administration	35%	57%	5%	9%	62%	19%	10%	10%	73%	12%	6%
Arts	24%	50%	**13%**	6%	57%	21%	**17%**	13%	64%	13%	**11%**
Lettres	29%	68%	2%	11%	56%	19%	14%	14%	64%	14%	8%
Échantillon total:	**36%**	**56%**	**5%**	**14%**	**59%**	**16%**	**10%**	**18%**	**65%**	**11%**	**6%**

Note: Les "ne sais pas" ne paraissent pas s'ils sont inférieurs à 5%. Les calculs sont basés sur les fréquences brutes. Se référer au chapitre 5, tableau C3-a, pour y trouver les proportions de l'échantillon total sur tous les points de l'échelle de réponses. La colonne "désaccord" inclut le "très en désaccord".

Tableau C3-b: **(suite) Critères de correction et représentativité**

	Q33i) Nécessaire d'établir les critères avant la correction N = 628 p = ,0021			Q33j) Les notes devraient refléter les objectifs maîtrisés N = 625 p = ,0029			
	TA	A	D	TA	A	D	NSP
Sc. biologiques	35%	56%	9%	32%	59%	6%	3%
Éduc. physique	50%	47%	----	23%	67%	3%	7%
Tech. biologiques	55%	43%	2%	26%	66%	1%	7%
Sc. physiques	31%	62%	3%	11%	72%	13%	2%
Tech. physiques	45%	53%	2%	28%	56%	11%	5%
Sc. humaines	40%	53%	5%	9%	57%	15%	19%
Tech. humaines	69%	30%	2%	33%	57%	8%	2%
Administration	54%	38%	7%	32%	56%	9%	4%
Arts	64%	31%	5%	25%	67%	4%	5%
Lettres	43%	51%	3%	22%	65%	10%	5%
Échantillon total:	**49%**	**46%**	**4%**	**24%**	**62%**	**8%**	**5%**

Note: Les "ne sais pas" ne paraissent pas s'ils sont inférieurs à 5%. Les calculs sont basés sur les fréquences brutes. Se référer au chapitre 5, tableau C3-a, pour y trouver les proportions de l'échantillon total sur tous les points de l'échelle de réponse. La colonne "désaccord" inclut le "très en désaccord".

Tableau P3-b: Les pratiques relatives aux qualités de l'évaluation

	Q15) Je prépare la plupart de mes questions d'examen à partir de questions déjà existantes N = 566 p = ,0015					Q13) Je fais examiner mes questions par un ou des collègues avant d'administrer un examen N = 568 p = ,0001					Q31) Je fais examiner mes consignes de travaux écrits par un ou des collègues avant de les présenter aux étudiants N = 452 p = ,0002				
	T	S	P	R	J	T	S	P	R	J	T	S	P	R	J
Sc. biologiques	9%	42%	27%	15%	6%	3%	21%	21%	24%	30%	4%	42%	8%	17%	29%
Éduc. physique	4%	41%	44%	11%	----	4%	32%	29%	18%	18%	11%	21%	26%	21%	21%
Tech. biologiques	4%	38%	39%	17%	3%	19%	33%	28%	11%	8%	18%	39%	22%	10%	12%
Sc. physiques	8%	51%	31%	8%	1%	7%	21%	32%	22%	18%	16%	16%	20%	25%	23%
Tech. physiques	----	40%	44%	11%	5%	2%	5%	24%	35%	35%	5%	5%	32%	27%	32%
Sc. humaines	4%	31%	46%	12%	6%	----	12%	31%	22%	35%	4%	28%	32%	17%	19%
Tech. humaines	4%	24%	35%	31%	6%	2%	14%	31%	33%	20%	2%	20%	33%	31%	14%
Administration	2%	37%	42%	17%	2%	1%	14%	23%	42%	21%	1%	17%	32%	33%	17%
Arts	3%	17%	46%	26%	9%	3%	11%	23%	40%	23%	6%	15%	38%	21%	21%
Lettres	2%	24%	29%	31%	14%	----	6%	30%	32%	32%	5%	10%	29%	36%	20%
Échantillon total:	**4%**	**36%**	**38%**	**17%**	**5%**	**5%**	**17%**	**27%**	**29%**	**23%**	**7%**	**20%**	**28%**	**25%**	**20%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau P3-b: **Pratiques relatives à la fréquence de l'évaluation**

	Q11) Dans ce cours, j'évalue les apprentissages, pour des fins de notation... N = 631 p = ,0001			
	à toutes les semaines	aux 2 semaines	aux 3 semaines	aux 5 semaines ou plus rarement
Sc. biologiques	21%	18%	44%	18%
Éduc. physique	17%	----	30%	53%
Tech. biologiques	7%	13%	49%	32%
Sc. physiques	14%	18%	55%	13%
Tech. physiques	14%	20%	44%	22%
Sc. humaines	16%	25%	47%	11%
Tech. humaines	13%	13%	44%	30%
Administration	15%	23%	41%	22%
Arts	22%	10%	49%	19%
Lettres	23%	26%	47%	5%
Échantillon total:	**16%**	**18%**	**46%**	**21%**

Tableau P3-b: **Pratiques de présentation de questions d'examen. Strates affichant des proportions de 3% et plus sur les réponses "Toujours", "Souvent" ou "Parfois".**

	Q20a) J'écris les questions au tableau ou je les projette sur acétates		Q20c) J'en fais la lecture à haute voix et les étudiants les prennent en note		
	Souvent	Parfois	Toujours	Souvent	Parfois
Éduc. physique	----	----	----	8%	4%
Sc. humaines	----	8%	3%	3%	3%
Administration	----	1%	5%	----	----
Arts	4%	8%	12%	----	----
Lettres	14%	5%	3%	3%	13%
Échantillon total:	**2%**	**3%**	**3%**	**2%**	**2%**

Note: N = 460; p = ,0001

Tableau P3-b: (suite) Administration de l'évaluation: "Dans mes examens à développement..."

	Q24) Je propose des questions "au choix" N = 477					Q25) Je permets aux étudiants de répondre "à la maison" N = 352				Q26) Je permets toute documentation utile N = 472				
	T	S	P	R	J	T+S	P	R	J	T	S	P	R	J
Sc. biologiques	----	26%	29%	16%	29%	----	4%	11%	86%	4%	4%	25%	18%	50%
Éduc. physique	----	----	26%	13%	**61%**	5%	14%	10%	71%	9%	14%	----	18%	59%
Tech. biologiques	----	4%	19%	19%	58%	----	8%	4%	88%	2%	2%	19%	15%	**62%**
Sc. physiques	----	6%	10%	**41%**	44%	----	3%	6%	**91%**	3%	9%	21%	**22%**	46%
Tech. physiques	----	4%	9%	36%	51%	----	4%	8%	88%	22%	18%	34%	14%	12%
Sc. humaines	**15%**	20%	30%	17%	17%	2%	**17%**	17%	64%	11%	19%	26%	11%	34%
Tech. humaines	4%	18%	24%	18%	36%	**7%**	17%	26%	50%	4%	15%	**41%**	15%	24%
Administration	1%	11%	16%	25%	46%	1%	6%	10%	83%	9%	20%	29%	18%	24%
Arts	8%	8%	**35%**	19%	31%	4%	16%	8%	72%	19%	22%	19%	15%	26%
Lettres	5%	**33%**	29%	12%	21%	----	15%	**30%**	52%	**28%**	**26%**	26%	9%	12%
Échantillon total:	**3%**	**12%**	**20%**	**24%**	**40%**	**2%**	**9%**	**13%**	**76%**	**10%**	**15%**	**26%**	**16%**	**33%**

Note: T = TOUJOURS; S = SOUVENT; P = PARFOIS; R = RAREMENT; J = JAMAIS.
Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total (p = ,0001).

Tableau P3-b: (suite) Pratiques de préparation des étudiants

	Q19a) Fais faire des exercices avec questions équivalentes N = 548 p = ,0001					Q19c) Fais des démonstrations ou des révisions N = 559 p = ,0033					Q32a) Explique la tâche N = 427 p = ,0339		Q32b) Feuille explicative N = 445 p = ,0265	
	T	S	P	R	J	T	S	P	R	J	R	J	P	R+J
Sc. biologiques	38%	28%	12%	9%	12%	22%	22%	28%	**19%**	9%	----	----	8%	8%
Éduc. physique	19%	19%	22%	19%	**22%**	30%	48%	19%	----	4%	----	----	**26%**	**21%**
Tech. biologiques	**1%**	31%	37%	16%	14%	7%	40%	26%	13%	**14%**	2%	----	8%	10%
Sc. physiques	**42%**	34%	12%	8%	4%	23%	34%	29%	8%	6%	5%	2%	7%	5%
Tech. physiques	10%	41%	30%	16%	3%	11%	29%	**37%**	13%	10%	**6%**	**9%**	12%	7%
Sc. humaines	21%	**44%**	10%	10%	15%	20%	43%	16%	14%	6%	2%	2%	13%	4%
Tech. humaines	11%	39%	27%	9%	14%	25%	46%	27%	2%	----	----	----	12%	2%
Administration	13%	**44%**	31%	7%	5%	21%	49%	20%	8%	3%	----	2%	14%	2%
Arts	10%	29%	26%	**26%**	10%	29%	21%	35%	12%	3%	----	6%	10%	13%
Lettres	20%	42%	16%	14%	8%	**31%**	41%	18%	6%	4%	----	----	20%	13%
Échantillon total:	**18%**	**37%**	**23%**	**12%**	**9%**	**21%**	**38%**	**25%**	**9%**	**6%**	**1%**	**2%**	**13%**	**7%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau P3-b: (suite) Pratiques de correction et de notation

	Q7) Lorsque je corrige, les critères d'évaluation sont déterminés... N = 625 p = ,0001				Q9) Je donne les critères de correction aux étudiants à l'avance N = 633 p = ,0001					Q27) Compensation du choix au hasard (guessing) dans les questions à choix multiples N = 395 p = ,0238			
	À l'avance	Juste avant	Au cours correction	Après 1ère lecture	T	S	P	R	J	T+S	P	R	J
Sc. biologiques	85%	9%	3%	3%	26%	32%	32%	9%	----	----	7%	18%	75%
Éduc. physique	94%	3%	3%	----	55%	29%	10%	6%	----	12%	4%	24%	60%
Tech. biologiques	95%	1%	----	4%	45%	40%	6%	8%	1%	7%	12%	8%	74%
Sc. physiques	70%	24%	3%	2%	23%	36%	22%	10%	9%	34%	9%	9%	48%
Tech. physiques	81%	14%	2%	3%	38%	30%	20%	11%	2%	12%	22%	20%	46%
Sc. humaines	66%	15%	2%	17%	46%	33%	11%	6%	4%	15%	12%	9%	64%
Tech. humaines	87%	4%	2%	8%	65%	26%	4%	6%	----	17%	11%	6%	66%
Administration	78%	15%	2%	5%	33%	36%	17%	12%	2%	7%	9%	16%	67%
Arts	92%	2%	5%	2%	58%	27%	8%	2%	5%	13%	13%	7%	67%
Lettres	90%	6%	----	3%	41%	44%	9%	3%	3%	18%	9%	17%	57%
Échantillon total:	**83%**	**11%**	**2%**	**5%**	**41%**	**34%**	**14%**	**8%**	**3%**	**12%**	**11%**	**13%**	**63%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau P3-b: Pratiques d'amélioration a posteriori

	Q14) Après la correction d'un examen, je fais l'analyse de mes questions d'examen N = 567 p = ,003			Q22) Demande opinion sur la formulation des items N = 568 p = ,0169		Q29) Je conserve mes questions d'examen N = 557 p = ,0008	
	T+S	P	R+J	R	J	P	R+J
Sc. biologiques	61%	33%	6%	**39%**	6%	9%	3%
Éduc. physique	68%	14%	**18%**	30%	19%	----	4%
Tech. biologiques	**90%**	10%	----	7%	14%	11%	1%
Sc. physiques	67%	21%	13%	31%	25%	24%	9%
Tech. physiques	58%	31%	11%	32%	17%	23%	7%
Sc. humaines	58%	29%	12%	22%	22%	19%	2%
Tech. humaines	69%	29%	2%	37%	8%	12%	10%
Administration	64%	29%	7%	29%	14%	19%	7%
Arts	57%	29%	**14%**	34%	14%	29%	11%
Lettres	61%	33%	6%	35%	**27%**	**31%**	**14%**
Échantillon total:	**66%**	**25%**	**8%**	**29%**	**17%**	**19%**	**7%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau C4-b: **Les croyances relatives aux instruments d'évaluation**

	Q33k) Les questions à développement peuvent seules mesurer les apprentissages de haut niveau N = 621 p = ,0018					Q33r) Les questions à choix multiples peuvent mesurer des apprentissages de haut niveau N = 621 p = ,0001					Q33o) Les questions à choix multiples mesurent principalement les connaissances N = 624 p = ,0001					Q33s) Les questions à choix multiples font l'objet de réponses au hasard trop souvent N = 623 p = ,0001			
	TA	A	D	TD	NSP	TA	A	D	TD	NSP	TA	A	D	TD	NSP	TA	A	D	NSP
Sc. biologiques	9%	**41%**	41%	3%	6%	3%	53%	32%	3%	9%	6%	44%	41%	9%	----	3%	32%	50%	15%
Éduc. physique	10%	24%	52%	7%	7%	----	43%	33%	7%	17%	----	**47%**	43%	7%	3%	3%	43%	33%	20%
Tech. biologiques	9%	17%	**61%**	6%	6%	**12%**	**59%**	20%	3%	7%	3%	17%	**64%**	12%	4%	6%	18%	**58%**	18%
Sc. physiques	9%	39%	36%	6%	**10%**	----	31%	37%	18%	14%	1%	26%	44%	**13%**	16%	7%	41%	26%	27%
Tech. physiques	2%	30%	55%	8%	6%	3%	40%	35%	16%	6%	5%	31%	47%	11%	6%	8%	43%	41%	8%
Sc. humaines	15%	38%	35%	5%	7%	----	40%	29%	**24%**	7%	4%	40%	36%	2%	**18%**	7%	27%	38%	27%
Tech. humaines	10%	33%	41%	12%	4%	2%	30%	35%	22%	11%	**11%**	42%	34%	9%	4%	11%	37%	33%	19%
Administration	11%	32%	37%	**14%**	7%	3%	35%	38%	16%	8%	9%	34%	46%	8%	4%	10%	38%	42%	11%
Arts	12%	30%	46%	5%	7%	----	30%	**51%**	8%	11%	9%	36%	45%	2%	9%	9%	**51%**	20%	20%
Lettres	**30%**	39%	22%	5%	5%	6%	22%	32%	23%	17%	**11%**	46%	23%	9%	11%	**18%**	40%	14%	28%
Échantillon total:	**12%**	**32%**	**42%**	**8%**	**7%**	**3%**	**37%**	**34%**	**15%**	**10%**	**6%**	**34%**	**43%**	**8%**	**8%**	**9%**	**37%**	**33%**	**19%**
																Note: D inclut TD			

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau C4-b: (suite) Les croyances relatives aux instruments d'évaluation

	Q33l) Examens devraient être à livre ouvert N = 616 p = ,0001					Q33q) Travaux et projets sont meilleurs moyens d'évaluation N = 624 p = ,0001					Q33m) Évaluations formatives devraient être auto-évaluations N = 619 p = ,0001				
	TA	A	D	TD	NSP	TA	A	D	TD	NSP	TA	A	D	TD	NSP
Sc. biologiques	3%	9%	53%	**24%**	12%	6%	15%	**62%**	12%	6%	15%	**44%**	26%	----	15%
Éduc. physique	**13%**	13%	47%	7%	20%	3%	20%	53%	7%	17%	----	20%	50%	**20%**	10%
Tech. biologiques	----	4%	**69%**	13%	14%	----	18%	57%	4%	21%	4%	35%	49%	8%	4%
Sc. physiques	2%	8%	63%	20%	7%	----	11%	**62%**	7%	20%	2%	35%	35%	16%	11%
Tech. physiques	6%	**32%**	48%	6%	6%	8%	34%	45%	5%	8%	3%	28%	47%	12%	9%
Sc. humaines	4%	15%	53%	9%	20%	9%	37%	31%	4%	19%	4%	11%	**54%**	17%	15%
Tech. humaines	4%	17%	57%	13%	9%	4%	**46%**	37%	2%	11%	4%	33%	50%	11%	2%
Administration	3%	21%	57%	10%	9%	7%	29%	51%	7%	7%	6%	28%	40%	17%	9%
Arts	4%	20%	43%	8%	25%	16%	**47%**	27%	5%	4%	5%	12%	**68%**	5%	9%
Lettres	8%	12%	47%	11%	22%	11%	**49%**	29%	3%	8%	3%	17%	52%	13%	14%
Échantillon total:	**4%**	**15%**	**55%**	**12%**	**13%**	**6%**	**30%**	**46%**	**5%**	**12%**	**4%**	**27%**	**47%**	**13%**	**9%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau P4-b: **Proportions de répondants qui disent utiliser l'examen ou le travail écrit pour effectuer l'évaluation de leurs étudiants**

	Q12) Utilisent l'examen N = 636	Q30) Utilisent le travail écrit N = 635
Sc. biologiques	94%	69%
Éduc. physique	90%	61%
Tech. biologiques	92%	64%
Sc. physiques	99%	50%
Tech. physiques	98%	64%
Sc. humaines	89%	85%
Tech. humaines	91%	96%
Administration	96%	77%
Arts	59%	56%
Lettres	76%	91%
Échantillon total:	**89%**	**71%**

Note: Les proportions inférieures à la moyenne de l'échantillon total sont mises en évidence (p = ,0001).

Tableau P4-b: **Dans mes examens, j'utilise les genres d'items suivants:**

	Q21a) Questions à développement court N = 557 p = ,0021				Q21b) Questions à développement long N = 523 p = ,0001					Q21c) Questions objectives, à choix multiples N = 554 p = ,0001				
	T	S	P	R+J	T	S	P	R	J	T	S	P	R	J
Sc. biologiques	**42%**	42%	15%	----	13%	23%	23%	23%	19%	36%	33%	6%	12%	12%
Éduc. physique	30%	44%	22%	4%	4%	20%	12%	28%	36%	33%	**56%**	----	7%	4%
Tech. biologiques	16%	48%	20%	16%	5%	6%	17%	21%	**51%**	**41%**	**52%**	4%	3%	----
Sc. physiques	35%	48%	13%	4%	**23%**	28%	20%	19%	10%	1%	6%	19%	**32%**	42%
Tech. physiques	17%	**75%**	8%	----	3%	14%	20%	**39%**	24%	2%	29%	31%	27%	12%
Sc. humaines	25%	56%	12%	6%	18%	24%	31%	20%	7%	10%	40%	8%	8%	33%
Tech. humaines	20%	59%	18%	2%	11%	26%	30%	22%	11%	10%	33%	16%	22%	18%
Administration	22%	67%	9%	2%	6%	19%	**40%**	19%	15%	13%	26%	24%	26%	12%
Arts	21%	59%	15%	6%	7%	30%	17%	27%	20%	6%	18%	**35%**	15%	26%
Lettres	18%	61%	20%	2%	11%	**43%**	28%	9%	11%	----	11%	17%	21%	**51%**
Échantillon total:	**24%**	**57%**	**14%**	**5%**	**11%**	**22%**	**26%**	**22%**	**20%**	**14%**	**28%**	**17%**	**19%**	**21%**

Note: Les proportions sont calculées sur les fréquences non pondérées, afin de faciliter les comparaisons.

Tableaux P4-b: **Distribution, dans l'ordre décroissant, de l'utilisation des diverses approches d'évaluation (Q34a à Q34q). Ici, "l'utilisation" est définie par la somme des proportions de répondants qui ont indiqué "toujours" et "souvent" à la question 34: "Utilisez-vous les approches suivantes pour effectuer l'évaluation de l'apprentissage de vos étudiants?"**

À partir de la moyenne observée dans l'échantillon total, les tableaux montrent les strates dont l'utilisation est supérieure ou égale à la moyenne.

Q34n) Examen écrit N = 638		Q34l) Exercices en classe N = 636		Q34a) Devoirs ou exercices à la maison N = 636	
Sc. physiques	99%	Arts	80%	Administration	79%
Sc. biologiques	94%	Administration	71%	Lettres	71%
Administration	94%	Lettres	76%	Tech. physiques	62%
Tech. physiques	86%	Éduc. physique	65%	Tech. humaines	61%
Tech. biologiques	85%			Arts	56%
Échantillon total:	**82%**	**Échantillon total:**	**61%**	**Échantillon total:**	**57%**

Tableaux P4-b: **(suite)**

Q34j) Résolution de problème N = 628		**Q34k) Démonstration de savoir-faire N = 623**		**Q34g) Travail écrit, recherche N = 636**	
Sc. physique	76%	Éduc. physique	80%	Lettres	76%
Administration	65%	Tech. biologiques	66%	Sc. humaines	70%
Arts	61%	Arts	66%	Tech. humaines	67%
Tech. physiques	59%	Administration	42%	Arts	47%
		Lettres	42%		
Échantillon total:	**49%**	**Échantillon total:**	**43%**	**Échantillon total:**	**39%**

Q34f) Rapport ou compte-rendu d'activités ou de laboratoire N = 633		**Q34h) Questions orales N = 630**		**Q34c) Étude de cas N = 632**	
Tech. physiques	73%	Tech. biologiques	47%	Administration	47%
Sc. biologiques	60%	Lettres	38%	Tech. biologiques	46%
Tech. biologiques	45%	Arts	30%	Tech. humaines	31%
Sc. physiques	44%			Arts	31%
Tech. humaines	39%				
Échantillon total:	**37%**	**Échantillon total:**	**25%**	**Échantillon total:**	**24%**

Tableaux P4-b: (suite)

Q34e) Mini-test ou jeux questionnaire N = 634 p = ,1394		Q34b) Dossier anecdotique, journal de bord N = 633		Q34i) Projet thématique N = 627	
Sc. biologiques	34%	Tech. biologiques	40%	Arts	48%
Tech. humianes	30%	Tech. humaines	35%	Tech. humaines	36%
Tech. biologiques	27%	Éduc. physique	32%	Lettres	17%
Sc. physiques	27%	Arts	24%	Sc. humaines	15%
Éduc. physique	26%				
Administration	24%				
Échantillon total:	**23%**	**Échantillon total:**	**17%**	**Échantillon total:**	**15%**

Q34d) Observation directe avec grille d'appréciation et/ou liste de vérification (check-list) N = 629		Q34o) L'autoévaluation N = 634		Q34q) Schéma ou réseaux de concepts N = 612 p = ,0977	
Éduc. physique	43%	Tech. biologiques	40%	Sc. biologiques	17%
Arts	31%	Éduc. physique	29%	Éduc. physique	13%
Tech. biologiques	29%	Tech. humaines	17%	Sc. humaines	13%
Tech. humaines	20%	Sc. biologiques	14%	Arts	13%
		Arts	12%	Tech. physiques	12%
				Tech. biologiques	10%
				Tech. humaines	7%
Échantillon total:	**15%**	**Échantillon total:**	**12%**	**Échantillon total:**	**8%**

Tableaux P4-b: **(suite)**

Q34p)	**L'évaluation par les pairs N = 629**
Éduc. physique	16%
Tech. humaines	11%
Arts	10%
Lettres	6%
Tech. biologiques	5%
Sc. humaines	4%
Échantillon total:	**5%**

Note: L'approche 34m, examen surprise, n'est pas présentée en tableau ci-haut n'étant utilisée que par 3% des répondants en moyenne.

Note: Pour tous ces tableaux, le χ^2 est significatif ($p < .05$) sauf pour Q34q), "schéma ou réseaux de concepts" ($p = .0977$) et pour Q34e), "mini-test ou jeux questionnaire" ($p = .1394$). Ici, les % de l'échantillon total sont calculés sur les fréquences non pondérées, pour fin de comparaisons.

Tableau C5-b: **Les croyances relatives aux dimensions interpersonnelles de l'évaluation**

	Q44a) Les plus travaillants méritent notes plus élevées N = 627 p = ,0467			Q44d) Dans la correction, on est influencé par l'identité du répondant N = 562 p = ,0446			Q44e) Conception de l'enseignement influence choix de stratégies en évaluation N = 588 p = ,1149		
	TA	A	D	A	D	TD	TA	A	D
Sc. biologiques	27%	37%	37%	38%	50%	12%	12%	69%	19%
Éduc. physique	21%	62%	17%	70%	30%	----	35%	62%	4%
Tech. biologiques	15%	32%	53%	56%	35%	9%	28%	60%	12%
Sc. physiques	15%	34%	51%	31%	56%	12%	22%	71%	7%
Tech. physiques	17%	43%	40%	53%	43%	4%	16%	72%	12%
Sc. humaines	13%	31%	56%	58%	31%	11%	41%	53%	6%
Tech. humaines	13%	48%	38%	50%	46%	4%	28%	68%	4%
Administration	10%	41%	49%	43%	42%	15%	27%	60%	14%
Arts	25%	42%	32%	57%	38%	5%	36%	51%	13%
Lettres	21%	39%	40%	46%	42%	12%	26%	60%	14%
Échantillon total:	**17%**	**40%**	**44%**	**48%**	**42%**	**9%** NSP: 12%	**27%**	**62%**	**11%** NSP: 7%

Note: Pour Q44a et Q44e, les réponses "désaccord" et "très en désaccord" sont combinées. Pour Q44d, les réponses "très en accord" et "en accord" sont combinées.

Tableau P5-b: **Analyse, strate par strate, des interventions sur des notes finales des étudiants. Pourcentages intra-strates de réponses affirmatives aux questions 45 et 46. Présentation dans l'ordre décroissant des pratiques. "Il m'arrive de..."**

	Échantillon total	101	109	199	201	299	300	399	400	500	600
Q45) Diminuer pour pénaliser:											
a) absences non motivées N = 633	18%	20%	52%	16%	5%	19%	18%	19%	8%	40%	17%
b) peu d'effort fourni N = 628	15%	3%	40%	9%	1%	18%	11%	15%	9%	40%	23%
Q46) Augmenter pour récompenser:											
e) progrès réalisés N = 630; p = ,0004	42%	46%	70%	29%	38%	47%	45%	25%	37%	57%	48%
a) motivation et effort N = 630; p = ,0107	41%	43%	60%	27%	38%	52%	38%	30%	42%	55%	44%
g) créativité, l'originalité N = 627	32%	23%	38%	19%	21%	34%	44%	30%	29%	58%	40%
f) participation en classe N = 624	31%	20%	70%	23%	20%	34%	29%	28%	30%	51%	35%
d) réalisation de travaux supplémentaires N = 621	22%	9%	14%	7%	15%	39%	28%	22%	17%	26%	34%
c) devoirs à la maison N = 624; p = ,0693	21%	11%	21%	13%	17%	31%	17%	23%	24%	34%	22%
b) présence assidue N = 625	19%	11%	55%	11%	9%	27%	16%	15%	18%	31%	20%

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total. Les degrés de signification sont indiqués lorsque différents de ,0001. Pour Q46c), le test "chi carré" n'est pas significatif (p = ,0693).

Tableau P5-b: **(suite) Pratiques d'intervention sur les notes versées au bulletin, favorisant soit le groupe-cours (Q23), soit des individus (Q47)**

	Q23) Je donne un examen plutôt facile pour compenser un examen antérieur difficile N = 565; p = ,0151;				Q47) J'encourage un étudiant faible en accordant une note de passage lorsque ses efforts me semblent appréciables N = 634 ; p = ,0001;				
	T+S	P	R	J	T	S	P	R	J
Sc. biologiques	3%	39%	30%	27%	----	6%	34%	**49%**	11%
Éduc. physique	----	11%	41%	**48%**	**19%**	**35%**	23%	19%	3%
Tech. biologiques	1%	27%	35%	37%	----	9%	37%	24%	**29%**
Sc. physiques	5%	37%	40%	18%	1%	9%	37%	34%	18%
Tech. physiques	6%	**46%**	32%	16%	3%	21%	37%	22%	17%
Sc. humaines	**10%**	29%	33%	29%	----	26%	37%	31%	6%
Tech. humaines	2%	27%	31%	41%	----	7%	31%	43%	19%
Administration	4%	32%	44%	21%	----	12%	38%	33%	17%
Arts	3%	29%	46%	23%	3%	**31%**	31%	28%	7%
Lettres	8%	36%	42%	14%	3%	20%	35%	32%	9%
Échantillon total:	**5%**	**32%**	**38%**	**25%**	**2%**	**16%**	**35%**	**31%**	**15%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau C6-b: **Analyse des croyances pour lesquelles il y a différence significative entre les strates**

	Q49e) Les enseignants devraient généralement commenter par écrit les productions écrites de leurs étudiants N = 607 p = ,001			Q49g) Les résultats d'une évaluation formative ne doivent pas servir au calcul de la note inscrite au bulletin N = 582 p = ,0235				Q49d) Plusieurs enseignants accordent une note de passage à des étudiants qui ne la méritent pas N = 375 p = ,0022			
	TA	A	D	TA	A	D	TD	TA	A	D	TD
Sc. biologiques	30%	64%	6%	21%	**55%**	21%	3%	----	19%	**56%**	**25%**
Éduc. physique	18%	**82%**	----	21%	32%	39%	7%	----	29%	**53%**	18%
Tech. biologiques	36%	64%	----	23%	49%	26%	1%	----	37%	**53%**	9%
Sc. physiques	21%	73%	6%	12%	52%	28%	7%	9%	33%	**52%**	6%
Tech. physiques	24%	63%	**13%**	20%	27%	**42%**	**10%**	8%	40%	45%	8%
Sc. humaines	38%	55%	8%	16%	38%	38%	9%	7%	**66%**	28%	----
Tech. humaines	**54%**	44%	2%	**32%**	32%	28%	8%	7%	46%	43%	4%
Administration	36%	62%	2%	17%	36%	38%	9%	7%	**62%**	22%	9%
Arts	31%	64%	5%	10%	25%	**52%**	**13%**	**11%**	57%	27%	5%
Lettres	48%	42%	**9%**	17%	39%	37%	7%	**10%**	55%	36%	----
Échantillon total:	**34%**	**61%**	**5%**	**19%**	**39%**	**35%**	**8%**	**7%**	**47%**	**39%**	**7%**

Note: Les calculs ne contiennent pas les "ne sais pas" et sont basés sur les fréquences brutes. C'est ce qui peut expliquer certaines différences entre ces tableaux d'analyse par strates et les tableaux présentant les fréquences pondérées des répondants dans leur ensemble. Se référer au chapitre 5, tableau C6-a, pour y trouver les proportions de "ne sais pas" ainsi que les proportions de l'échantillon total.

Tableau P6-b: **Les pratiques de rétroaction et de notation: Analyse stratifiée**

	Q 28) Travail en équipe: même note à tous les membres de l'équipe N = 449 p = ,0092					Q 50) Note finale contient des notes obtenues en évaluation formative N = 634 p = ,0001					Q 52) Rétroaction corrective N = 635 p = ,0079		Q53) Remédiation N = 633 p = ,176	
	T	S	P	R	J	T	S	P	R	J	P	R+J	P	R+J
Sc. biologiques	62%	23%	12%	----	4%	6%	3%	9%	29%	53%	20%	**14%**	**26%**	**9%**
Éduc. physique	42%	37%	16%	5%	----	7%	23%	23%	20%	27%	30%	**20%**	20%	3%
Tech. biologiques	43%	36%	15%	4%	2%	12%	14%	4%	12%	**59%**	12%	5%	9%	3%
Sc. physiques	**64%**	20%	7%	2%	7%	11%	11%	14%	12%	51%	**36%**	10%	23%	2%
Tech. physiques	25%	**50%**	12%	9%	4%	14%	**32%**	14%	10%	30%	33%	8%	**27%**	2%
Sc. humaines	49%	46%	5%	----	----	**20%**	9%	9%	11%	50%	29%	9%	20%	4%
Tech. humaines	33%	33%	19%	7%	7%	11%	17%	15%	13%	44%	19%	2%	17%	4%
Administration	30%	45%	10%	**12%**	3%	15%	18%	8%	24%	36%	24%	**14%**	21%	6%
Arts	43%	25%	14%	11%	7%	**28%**	23%	16%	10%	23%	20%	8%	13%	----
Lettres	30%	38%	12%	8%	**12%**	18%	21%	17%	6%	38%	18%	3%	15%	----
Échantillon total:	**40%**	**37%**	**12%**	**6%**	**5%**	**15%**	**17%**	**12%**	**14%**	**42%**	**24%**	**9%**	**19%**	**3%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

FORMATION ET PERFECTIONNEMENT

Tableau Q55: **Analyse stratifiée des sources de formation en évaluation des apprentissages (Q55) N = 628 ...en ordre décroissant**

	Q55-3) Au moins un cours CPEC-Performa p = ,0001	Q55-6) Au moins une conférence ou un atelier p = ,0085	Q55-2) Au moins un cours au bacc. p = ,0001	Q55-4) Au moins un cours de maîtrise p = ,0006	Q55-1) Aucune p = ,0269
Sc. biologiques	49%	40%	14%	9%	14%
Éduc. physique	29%	32%	**77%**	13%	3%
Tech. biologiques	**60%**	42%	36%	**32%**	5%
Sc. physiques	36%	36%	28%	8%	14%
Tech. physiques	33%	17%	31%	14%	16%
Sc. humaines	31%	31%	22%	15%	**24%**
Tech. humaines	46%	35%	26%	17%	11%
Administration	42%	40%	23%	12%	11%
Arts	38%	34%	33%	5%	16%
Lettres	18%	**55%**	45%	18%	5%
Échantillon total:	**39%**	**37%**	**32%**	**15%**	**12%**

Notes: Les sources de formation en évaluation sont placées, ici, dans l'ordre décroissant des proportions observées pour l'ensemble des répondants. Les deux sources de formation les plus "fréquentées" n'apparaissent pas ici parce qu'il n'y a pas de différence significative entre les strates. Ces deux sources sont les journées pédagogiques (40%) et l'autodidaxie (39%).
Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau Q56: **Autoévaluation des compétences en évaluation des apprentissages. "Comment jugez-vous vos compétences dans les domaines suivants?"**

	Q56a) Évaluation formative N = 662; p = ,0203				Q56b) Évaluation sommative N = 665; p = ,0242			
	TB	B	P	F ou N	TB	B	P	F
Sc. biologiques	6%	40%	34%	20%	26%	66%	9%	----
Éduc. physique	13%	57%	23%	7%	19%	74%	3%	3%
Tech. biologiques	13%	65%	18%	4%	15%	72%	12%	1%
Sc. physiques	9%	55%	24%	11%	27%	69%	3%	----
Tech. physiques	16%	50%	22%	12%	16%	72%	8%	5%
Sc. humaines	2%	50%	27%	21%	13%	73%	8%	6%
Tech. humaines	17%	44%	28%	11%	11%	57%	26%	6%
Administration	16%	51%	25%	8%	22%	69%	8%	1%
Arts	25%	51%	25%	----	25%	54%	18%	4%
Lettres	9%	55%	26%	9%	25%	63%	9%	3%
Échantillon total :	**13%**	**53%**	**25%**	**10%**	**20%**	**67%**	**10%**	**2%**

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total.

Tableau Q66: **Proportions de répondants ayant indiqué qu'ils utilisent l'une ou plusieurs des sept formules pédagogiques ci-après. Ces formules pédagogiques sont présentées dans l'ordre décroissant des proportions observées dans l'échantillon**

	1- L'exposé magistral formel	2- Le travail individuel (exercices individuels)	3- Le travail en équipe	4- Le laboratoire	5- L'exposé informel en classe	6- Le séminaire ou groupe de discussion	7- Le stage
Sc. biologiques	71%	66%	69%	**91%**	51%	9%	----
Éduc. physique	39%	81%	**90%**	81%	61%	16%	3%
Tech. biologiques	69%	55%	74%	65%	42%	28%	**54%**
Sc. physiques	85%	75%	52%	53%	41%	1%	----
Tech. physiques	77%	69%	66%	81%	30%	6%	5%
Sc. humaines	84%	55%	71%	20%	49%	29%	----
Tech. humaines	76%	65%	**83%**	70%	**67%**	**57%**	**31%**
Administration	**86%**	80%	76%	56%	41%	18%	6%
Arts	57%	**84%**	61%	59%	59%	36%	18%
Lettres	62%	**85%**	82%	23%	**67%**	39%	2%
Échantillon total:	**74%**	**72%**	**71%**	**57%**	**49%**	**23%**	**13%**

N = 638; $p < .05$

Note: Proportions calculées sur fréquences non pondérées, afin de faciliter les comparaisons entre les strates et l'échantillon total. Puisque les répondants pouvaient cocher plusieurs items, les sommes horizontales excèdent 100%.

ANNEXE 6
ANALYSES FACTORIELLES

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
		*				
Q4	.66016	*	1	2.55700	19.7	19.7
Q5	.59168	*	2	1.67568	12.9	32.6
Q6A	.55295	*	3	1.28257	9.9	42.4
Q6B	.57155	*	4	1.04310	8.0	50.4
Q6C	.55613	*	5	1.01019	7.8	58.2
Q6D	.60466	*				
Q6E	.45104	*				
Q6F	.54206	*				
Q6G	.66051	*				
Q6H	.62057	*				
Q6I	.61529	*				
Q6J	.58516	*				
Q6K	.55677	*				

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Q6G	.80350				-.13505
Q6H	.72084	.22151	-.16408	.31648	.27974
Q6B	.59227	.20330			
		.71466	-.14339	.12497	
Q6C		.60855		-.37691	.26889
Q6D	-.14116	.55885		.11693	.14997
Q6E	.31975				.20709
		-.11459	.74131		
Q6I			.71294	-.12041	
Q6F		.24921	.51402	.21251	-.46828
Q5					
		.11691	-.13767	.79062	
Q4		-.27134	.34792	.48348	.38006
Q6J	-.11087				.69227
		.25634			
Q6K				.37240	.46948
Q6A	.43340				

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
		*				
Q2A	.62753	*	1	3.45458	15.7	15.7
Q2B	.59522	*	2	2.26811	10.3	26.0
Q2C	.36594	*	3	1.73794	7.9	33.9
Q2D	.49938	*	4	1.42970	6.5	40.4
Q2E	.68653	*	5	1.25978	5.7	46.1
Q2F	.64714	*	6	1.18258	5.4	51.5
Q2G	.60182	*	7	1.14110	5.2	56.7
Q3A	.62496	*	8	1.03086	4.7	61.4
Q3B	.54180	*				
Q3C	.55412	*				
Q3D	.47490	*				
Q3E	.64858	*				
Q3F	.57345	*				
Q3G	.57206	*				
Q6L	.65441	*				
Q6M	.77172	*				
Q6N	.74381	*				
Q6O	.70897	*				
Q6P	.70260	*				
Q6Q	.70373	*				
Q6R	.61213	*				
Q43	.59383	*				

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Q3E	.79339							
Q3A	.78122					-.20066		
Q3B	.69368							
Q3D	.65701	.11371						
Q2G	.52260	.33686	-.42797	.12290				.22270
Q2C	.35934	.14572	-.12905	.31557	.21084			
Q2E		.79985	.11758	.14716				-.10281
Q2F	.17744	.76959	-.12287					
Q2D		.63558		.21527	.14887			
Q3F	.23713	.53197	-.14428	-.41406			.18668	
Q6O			.80019			-.15502		.15894
Q6P	-.20142		.79512			.14384		
Q2A				.75148				-.20510
Q2B		.32756		.64707				.22504
Q6N			.14828		-.82140	.12162		.17494
Q6L			.18156	.15433	.69933	.10914		.27879
Q6Q		-.14751				.80782	.10478	
Q6R		.38476			.14614	.61435	-.20341	-.11420
Q3G		.22724	-.21512		.16417		.65854	
Q43		-.12229		.20975	-.18412	.33528	.60954	.12294
Q3C	.19925	.12643	.22926		.10871	-.15568	.59302	-.22739
Q6M		-.10835				-.10565		.85654

Final Statistics:

Variable	Communality	Factor	Eigenvalue	Pct of Var	Cum Pct
Q7	.64744	1	4.16316	11.9	11.9
Q8	.65839	2	2.72136	7.8	19.7
Q9	.51324	3	2.05350	5.9	25.5
Q10	.63474	4	1.86502	5.3	30.9
Q11	.61900	5	1.69655	4.8	35.7
Q13	.80476	6	1.65747	4.7	40.4
Q14	.61098	7	1.51331	4.3	44.8
Q15	.65597	8	1.36672	3.9	48.7
Q16	.70859	9	1.31377	3.8	52.4
Q17	.65736	10	1.22566	3.5	55.9
Q18	.58892	11	1.17116	3.3	59.3
Q19A	.47043	12	1.10592	3.2	62.4
Q19B	.71741	13	1.05933	3.0	65.5
Q19C	.54426				
Q20A	.85482				
Q20B	.88276				
Q20C	.54178				
Q22	.63397				
Q24	.54781				
Q25	.68632				
Q26	.61041				
Q27	.62737				
Q29	.70764				
Q31	.75172				
Q32C	.61096				
Q33A	.73451				
Q33B	.79542				
Q33C	.76678				
Q33D	.53490				
Q33E	.71814				
Q33F	.56253				
Q33G	.60066				
Q33H	.55336				
Q33I	.69771				
Q33J	.66183				

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Q7	.71895	.15324			.10219	-.13084	-.18413	
Q33I	.68089		.17946		-.11601		.17899	.31556
Q9	.61137	.11243				-.16436	.17582	
Q17	.57037	.19994		-.11203		.24929		
Q13		.86034	.20186					
Q31	.13772	.82236	.13802					
Q33B		.19728	.86024			.10991		
Q33C		.22563	.79694	.11986			.13850	.15274
Q33A	.18404		.71218					.10517
Q20B				.92773				
Q20A				-.90661				
Q25					.79614	.12757		
Q26	-.13064				.64196	-.25506		
Q24	.20480		-.12537	-.17663	.52525		.31466	.15974
Q15		.13659		-.11480		.76105		
Q29	-.14751	-.21367		.14706	-.17915	.71717		.10935
Q32C	-.12204				.11953	-.15712	.69304	
Q22	.24730	.28878	.19899				.58771	-.12940
Q19B	.12816	.22010				.20687	.53026	.19342
Q33D					-.14418			.69286
Q33F		-.15970			.22836			.67011
Q33G								.24512
Q33E	-.19245	.16793		.11474	.26033			.21001
Q19A	.16567	.13649		.14194	.26890		.21365	-.10662
Q18	.26804	.11263	-.10996	.12069	.19499	.27321	-.29649	.26396
Q19C			-.11211	.14466	.31343	-.16521	.29798	
Q27					.18428	.15119		
Q33J	.26383		.13749			-.20810		
Q8	.15518	.41079				-.17980		.21943
Q16							.18090	
Q14	.10687	.27156	.10917					
Q33H	.10299							.25445
Q10	-.12495	-.15650				.30812	-.12964	-.31559
Q20C		-.11054			.26344	-.20319		
Q11					.14364	-.13159		

	Factor 9	Factor 10	Factor 11	Factor 12	Factor 13
Q7				.15531	
Q33I		.20670			
Q9		-.15881			
Q17	.15852	.30584	.25630	-.11166	
Q13					
Q31					
Q33B					
Q33C		.14259			.14239
Q33A			.14526		-.33653
Q20B					
Q20A					
Q25				-.10838	
Q26		.31135			
Q24					.14836
Q15			-.12305		
Q29			.16666		
Q32C	.14505	.12197	.13456		
Q22		-.10187			-.22890
Q19B		.15794	-.15862	-.31510	.37225
Q33D	.12349				
Q33F					
Q33G	.71090				
Q33E	-.61647	-.22169			-.28113
Q19A	.49438			-.11053	
Q18	.33675		.17284	-.21361	-.12236
Q19C	.33307		.28047	.13051	.25087
Q27	.16007	.68259	-.17516		-.11712
Q33J	-.18242	.67293	.14952		
Q8		.44702	.37884	-.13867	
Q16			.79950		.21018
Q14		.13235	.53760	-.22331	-.31502
Q33H	.14048			.64316	-.14549
Q10			.11193	.59024	
Q20C	.19772			-.55167	-.23202
Q11	.11311		.13667		.73390

Note: Tel qu'expliqué aux pages 169 et 202, il n'y a pas d'analyse factorielle pour le quatrième champ de compétence.

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
		*				
Q44A	.67727	*	1	2.10973	21.1	21.1
Q44B	.67285	*	2	1.26098	12.6	33.7
Q44C	.51411	*	3	1.19774	12.0	45.7
Q44D	.46282	*	4	1.07644	10.8	56.4
Q44E	.65511	*				
Q47	.49346	*				
Q48	.32302	*				
Q23	.61794	*				
RECOMP	.60196	*				
PENAL	.62634	*				

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4
PENAL	.72036	.16680	.21670	-.18067
Q44A	-.71461	.20341	.27586	-.22166
RECOMP	.61843	.46078		
Q44C		.69008	.11734	.14905
Q48		-.49465		.25379
Q47	-.38175	-.48085	.32538	.10314
Q44B		.15595	.79865	
Q23	.14058	-.42949	.64278	
Q44E		.17037		.79101
Q44D		-.22495		.64002

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
		*				
Q28	.47728	*	1	1.88350	12.6	12.6
Q49A	.54026	*	2	1.72268	11.5	24.0
Q49B	.59793	*	3	1.53352	10.2	34.3
Q49C	.76798	*	4	1.22129	8.1	42.4
Q49D	.42710	*	5	1.17558	7.8	50.2
Q49E	.51839	*	6	1.08140	7.2	57.5
Q49F	.73891	*	7	1.04341	7.0	64.4
Q49G	.77779	*				
Q49H	.66778	*				
Q49I	.72965	*				
Q50	.80027	*				
Q51	.76442	*				
Q52	.71315	*				
Q53	.64657	*				
Q54	.49390	*				

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Q50	-.88495						
Q49G	.87292						
				-.18350		.18053	
Q52		.80180			.16254	-.36105	
Q53		.68896				.20896	
Q49E		.54428	-.15914	.36555			
			.76363				
Q49B			.56987	.11300	.23180	.14159	.31672
Q49A	-.15315		-.42080		.41728	-.26766	
Q28	.22977						
				.82992	-.14482		
Q49I				.58829	.35908	.13822	-.10441
Q49H		-.19366	.35225				
					.69671		
Q54					-.53037		
Q49D	-.19272	-.19169	-.15944	.21284			
						.82995	
Q49F	.15114		.12347				
					-.17634	-.25701	.71749
Q49C	.23958		.29246		-.21152	-.31988	-.70399
Q51	.20658	.19144	.20283				

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
Q44A	.57083	*	1	2.26118	14.1	14.1
Q44C	.43016	*	2	1.74122	10.9	25.0
Q33H	.54153	*	3	1.42924	8.9	33.9
Q49D	.55912	*	4	1.32150	8.3	42.2
Q47	.50963	*	5	1.28431	8.0	50.2
Q48	.50021	*	6	1.09222	6.8	57.1
Q44B	.68323	*				
Q23	.67989	*				
Q25	.58721	*				
Q49C	.54530	*				
Q6D	.53629	*				
Q6F	.69042	*				
Q6J	.54023	*				
Q6I	.53311	*				
RECOMP	.64915	*				
PENAL	.57333	*				

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
RECOMP	.79612					
PENAL	.72541				.18392	
Q47	-.52527	.29368	.29152	.17081		-.17331
Q6F		.74396	-.19520	.20562	.22964	
Q6I		.66738		-.25785		
Q6J		.63707	.15027	-.21449	-.25110	
Q23			.80348			-.14826
Q48	-.26891		.61003	-.12875		.15298
Q25		-.23871		.67653	-.18501	-.17250
Q6D	-.27168			.60292	.25361	.15410
Q44B	.16131		.46090	.54450	-.32570	.19209
Q49D			.20048		.70326	.13920
Q44A	-.34371		.17513		-.64088	
Q33H		.11567		.24690		.67062
Q49C		.29103		.30124		-.60612
Q44C	.38341		-.18795			.49484

ANNEXE 7

CATÉGORIES D'ÉNONCÉS DE CROYANCES ET DE PRATIQUES

Le texte qui suit constitue une adaptation libre du texte publié en 1991 par Stiggins sous le titre "Relevant classroom assessment training for teachers". Il propose une description des champs de compétence que devraient avoir les enseignants en matière de mesure et évaluation des apprentissages en salle de classe. Ces champs constituent une excellente structure, en six catégories, qui facilite l'analyse de la tâche des professeurs au chapitre des pratiques d'évaluation en salle de classe et qui guide la planification du perfectionnement. Ce texte met en valeur les "catégories de Stiggins" à cause de leur utilité tant en recherche qu'en formation des maîtres.

Ce texte qui a été utilisé, tel quel, lors de la pré-validation des énoncés de notre questionnaire de recherche auprès d'un groupe de consultants-experts (chapitre 4: Méthodologie). Dans cette étape, le texte était désigné sous l'expression "Document 3". Nous avons aussi présenté, ailleurs, les "catégories de Stiggins" à titre de typologie des champs de compétence en évaluation des apprentissages (Howe, 1992).

1. L'utilisation de l'évaluation en classe

Les enseignants se servent de l'évaluation pour trois catégories de raisons:

a. D'abord, ils évaluent les apprentissages afin d'appuyer des décisions. Par exemple, ils voudront diagnostiquer les besoins des étudiants, regrouper des étudiants selon certaines caractéristiques, accorder des notes, choisir des étudiants en vue d'événements ou de services spéciaux;

b. Ensuite, ils évaluent afin d'aider ou de guider l'enseignement. Ce faisant, ils feront savoir aux étudiants ce qui est attendu d'eux, fourniront des occasions d'exercices, impliqueront les étudiants dans l'autoévaluation et dans l'évaluation par les pairs afin de les aider à apprendre, etc.;

c. Enfin, les enseignants se servent de l'évaluation comme mécanisme de gestion de la classe, pour faire travailler les étudiants ou pour contrôler certains comportements. C'est ici qu'on dit de l'évaluation qu'elle est l'outil et l'expression du pouvoir du professeur sur les étudiants.

Indépendamment de toute discussion de pertinence quant à l'utilisation de l'évaluation comme moyen de contrôle ou comme outil d'enseignement, les enseignants devraient comprendre les incidences des diverses utilisations de l'évaluation sur la qualité de l'enseignement et de l'apprentissage. De même, pour pouvoir attester de la qualité des évaluations, les enseignants doivent être bien au fait des différentes variables qui entrent en jeu dans chacune des trois catégories d'utilisation de l'évaluation.

2. Les cibles de l'évaluation

S'il est nécessaire que les professeurs possèdent une excellente compréhension des objectifs pédagogiques que les étudiants doivent atteindre, ils doivent aussi connaître l'ensemble des moyens qui sont disponibles pour évaluer l'atteinte de ces objectifs. C'est là un principe fondamental de toute bonne évaluation.

Par exemple, un professeur ne peut pas évaluer le niveau d'habileté d'un étudiant en matière d'écriture si ce professeur n'a pas une vision claire et détaillée des attributs d'une bonne écriture.

Une analyse, même sommaire, des plans de cours, des travaux et exercices prévus par la plupart des professeurs montre que ces derniers visent au moins cinq grands types d'apprentissages:

1. la connaissance d'un contenu disciplinaire, les connaissances déclaratives;

2. les habiletés intellectuelles de niveau supérieur ou les habiletés de résolution de problème;

3. les savoir-faire de type comportemental;

4. les savoir-faire de type production;

5. les attitudes.

Plusieurs de ces apprentissages sont fort complexes et les méthodes appropriées à leur évaluation ne sont pas nécessairement maîtrisées par les professeurs. Pour enseigner et évaluer avec compétence, les professeurs doivent posséder une vision claire des résultats attendus de l'apprentissage. Mais puisqu'ils doivent aussi évaluer ces résultats d'apprentissage, il leur faut connaître et savoir utiliser les méthodes d'évaluation qui correspondent le mieux aux apprentissages ciblés.

Pour chacun des types d'apprentissages visés, les professeurs doivent savoir faire correspondre les dispositifs d'évaluation appropriés. Plusieurs cibles d'évaluation peuvent être mesurées par des examens objectifs. Plusieurs autres exigent des dispositifs fort différents.

3. Les qualités de l'évaluation

Les caractéristiques d'une "bonne" évaluation varient selon le contexte de l'évaluation. Il n'y a donc pas de formule, de recette, pour que toutes les évaluations en salle de classe soient menées avec les mêmes caractéristiques et les mêmes standards de qualité.

Cependant, les professeurs ont besoin de connaître et d'appliquer quelques critères communs à toutes les situations d'évaluation et de les adapter en fonction des contextes. Par exemple, nous savons qu'une évaluation de qualité.

- s'appuie sur un objet à évaluer, qui est explicite, spécifique, bien compris et auquel correspond une méthodologie, un dispositif approprié;
- s'assure de contrôler les variables qui deviennent des sources d'erreur de mesure. Ces variables sont présentes dans certaines caractéristiques des étudiants, dans le processus de mesure, dans l'environnement de mesure. Elles sont externes aux compétences mesurées mais elles peuvent influencer les résultats;
- mesure un échantillon des résultats de l'apprentissage qui soit représentatif des compétences ciblées et qui permette d'appuyer des conclusions, inférences ou généralisations. Ceci sans pour autant produire plus d'information qu'il n'en faut;
- fournit une information qui soit compréhensible et utilisable selon les buts recherchés.

Tous ces attributs d'une bonne évaluation contiennent diverses sources d'erreur de mesure. Les professeurs doivent savoir repérer et éviter ces sources d'erreur. Ils doivent pouvoir se rendre compte qu'un dispositif de mesure ne convient pas dans tel ou tel contexte, afin de corriger leur stratégie au besoin. Ils doivent contrôler et éliminer les sources d'interférence et de biais, tant dans l'évaluation dite objective que subjective, échantillonner correctement les résultats d'apprentissage, créer des dispositifs de mesure appropriés aux buts recherchés, et, enfin, s'assurer que les étudiants et les utilisateurs peuvent interpréter correctement l'information obtenue.

Ainsi, les qualités de l'évaluation ne s'expriment pas seulement en termes de fidélité et de validité des instruments de mesure. Les enseignants doivent pouvoir résoudre plusieurs problèmes relatifs à la conception, à l'administration, à la notation des dispositifs de mesure. De même, il faut pouvoir résoudre des problèmes relatifs à l'interprétation des notes et à l'utilisation qui sera faite des notes, notamment, en termes de conséquences sociales et psychologiques.

4. Les instruments d'évaluation

Les professeurs utilisent au moins trois formes d'évaluation des apprentissages en salle de classe. D'abord, ils utilisent des instruments de mesure de type "papier-crayon". Ce sont des questionnaires, des quiz, des descriptions de tâches à accomplir, la plupart du temps rédigés par les professeurs eux-mêmes. Ensuite, ils font appel à l'observation et à leur jugement expert dans le cadre de comportement ou de tâches, de la mise en oeuvre de procédures, par exemple au laboratoire, en stage, en séminaire. Troisièmement, les professeurs font de l'évaluation lors de communications personnelles directes avec les étudiants (entrevues, questions orales en salle de classe, conversations informelles) ou lors de discussions avec d'autres personnes. Dans cette dernière catégorie, le professeur fait souvent usage de son intuition, de son flair.

Toutes ces formes d'évaluation peuvent être correctement ou incorrectement utilisées. Chaque méthode a ses avantages et ses inconvénients et peut être plus ou moins appropriée à chaque contexte. Les professeurs doivent savoir s'en servir tout en comprenant que les règles de validité sont différentes d'une forme à l'autre.

5. Les dimensions interpersonnelles de l'évaluation

L'évaluation en salle de classe implique des jeux interpersonnels très complexes. Dans un tel environnement, l'évaluation est rarement scientifique, objective, désintéressée. Au contraire, elle est liée à toutes sortes de variables qui interviennent en amont comme en aval de la mesure proprement dite. Les professeurs d'expérience connaissent bien ces variables. Mais il n'est pas certain qu'ils comprennent en quoi elles jouent sur les dispositifs de mesure qu'ils choisissent d'utiliser ou encore sur leur jugement ou même sur l'interprétabilité des notes.

Les étudiants eux-mêmes sont à l'origine de plusieurs de ces variables personnelles:

- Ils arrivent en classe avec des antécédents culturels très variés, ayant dans certains cas des influences directes sur les résultats des évaluations;

- Ils ont des attentes et une image de soi héritées en bonne partie des évaluations des professeurs précédents;

- Ils utilisent eux-mêmes les résultats d'évaluation parce qu'ils s'autoévaluent et prennent des décisions en conséquence;

- Ils évaluent leurs chances d'atteindre les objectifs d'un cours;

- Ils évaluent leurs pairs, se jugent entre eux, ce qui contribue à la formation de certaines associations ou partenariats;

- Ils se forgent des conceptions très personnelles et très variées de ce qui peut constituer des critères d'excellence ou de ce qui appartient au domaine des objectifs implicites du cours;

- Ils sont très différents les uns des autres en termes de tempéraments, de besoins de feedback. Ils affichent divers degrés de motivation ou d'anxiété devant l'évaluation.

Quant à eux, les professeurs contribuent aussi au lot de variables personnelles qui jouent dans le processus d'évaluation. Ils sont naturellement très différents les uns des autres aux chapitres des:

- attentes à l'égard des étudiants;
- des attitudes envers les individus qui sont en classe;
- de leur conception de l'enseignement, de l'apprentissage, de l'évaluation, du contrôle de la salle de classe.

Comme les étudiants, ils sont très différents les uns des autres en termes de tempéraments, de motivations à enseigner et d'aptitudes à réagir aux personnes.

L'enseignant compétent doit savoir reconnaître et contrôler ces nombreuses dimensions personnelles s'il veut assurer à ses étudiants une évaluation équitable.

6. La rétroaction et la notation

Les enseignants fournissent continuellement des rétroactions à partir des résultats des évaluations. Toutes les rétroactions sont importantes mais celle qui exerce le plus d'influence et qui a les impacts les plus remarquables, c'est celle qui prend la forme d'une note qui sera versée au bulletin.

Les pratiques de notation sont plus appropriées lorsque les professeurs:

- Informent à l'avance les étudiants de leur politique et de leurs attitudes à ce sujet;
- Construisent la note finale autour des variables pertinentes (l'atteinte des objectifs, les résultats d'apprentissage) et ne la "contaminent" pas de variables non pertinentes (la participation en classe, la présence, la conduite, etc.);
- Introduisent dans la note finale des résultats fondés sur des évaluations de qualité;
- Maintiennent des données complètes et appropriées quand aux apprentissages de leurs étudiants;
- Pondèrent correctement les diverses notes accumulées pendant une session et établissent convenablement les seuils de réussite afin de traduire en note les compétences atteintes.

Quant aux autres formes de rétroaction, (la communication verbale et non verbale, les annotations sur les travaux écrits et compositions, les grilles d'évaluation sur échelles de jugement) les professeurs ont des pratiques appropriées lorsqu'ils s'assurent que leur rétroaction:

- correspond à ce qui est attendu explicitement dans l'objectif visé;
- est transmise dans le temps approprié et;
- est signifiante pour l'étudiant.

Code de diffusion: 1532-0412